

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: December 6, 2006, 10:48:46 ; Search time 75 Seconds  
(without alignments)  
138.397 Million cell updates/sec

Title: US-09-824-787B-2

Perfect score: 597

Sequence: 1 MSGEPQTSVAPPPPEVEPG.....ASNGETLEKITNSRPPCVIL 115

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 358988 seqs, 90258633 residues

Total number of hits satisfying chosen parameters: 358988

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications AA New.\*

- 1: /EMC\_Celerra\_SID83/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 2: /EMC\_Celerra\_SID83/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 3: /EMC\_Celerra\_SID83/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 4: /EMC\_Celerra\_SID83/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 5: /EMC\_Celerra\_SID83/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 6: /EMC\_Celerra\_SID83/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 7: /EMC\_Celerra\_SID83/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*
- 8: /EMC\_Celerra\_SID83/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	597	100.0	115	7	US-11-323-964-11
2	119	19.9	232	6	US-10-449-902-30557
3	104	17.4	257	6	US-10-449-902-33043
4	104	17.4	257	6	US-10-449-902-43851
5	101.5	17.0	87	6	US-10-533-519-233
6	93	15.6	212	6	US-10-953-349-16531
7	93	15.6	235	6	US-10-953-349-16530
8	93	15.6	236	6	US-10-953-349-16529
9	80	13.4	137	7	US-11-395-249-96
10	80	13.4	158	6	US-10-953-349-10235
11	80	13.4	158	7	US-11-056-355B-106087
12	80	13.4	158	7	US-11-056-355B-117326
13	80	13.4	176	6	US-10-953-349-15768
14	80	13.4	176	7	US-11-056-355B-53080
15	80	13.4	184	6	US-10-953-349-15767
16	80	13.4	184	7	US-11-056-355B-53079
17	80	13.4	186	6	US-10-953-349-10234
18	80	13.4	186	7	US-11-056-355B-106086
19	80	13.4	186	7	US-11-056-355B-117325
20	80	13.4	191	6	US-10-953-349-15766
21	80	13.4	191	7	US-11-056-355B-53078
22	80	13.4	203	7	US-11-056-355B-106085
23	80	13.4	203	7	US-11-056-355B-117324
24	76.5	12.8	170	6	US-10-953-349-25971
25	76.5	12.8	170	7	US-11-056-355B-63255

#### ALIGNMENTS

##### RESULT 1

US-11-323-964-11  
; Sequence 11, Application US/11323964  
; Publication No. US20060159689A1  
; GENERAL INFORMATION:  
; APPLICANT: Chiang, Chih-Sheng  
; APPLICANT: Simard, John J.L.  
; APPLICANT: Diamond, David C.  
; APPLICANT: Bot, Adrian Ion  
; APPLICANT: Liu, Xiping  
; TITLE OF INVENTION: COMBINATIONS OF TUMOR-ASSOCIATED  
; TITLE OF INVENTION: ANTIGENS IN DIAGNOSTICS FOR VARIOUS TYPES OF CANCERS  
; FILE REFERENCE: MANKK.050CP1  
; CURRENT APPLICATION NUMBER: US/11/323,964  
; CURRENT FILING DATE: 2005-12-29  
; PRIOR APPLICATION NUMBER: 11/155,288  
; PRIOR FILING DATE: 2005-06-17  
; PRIOR APPLICATION NUMBER: PCT/US2005/021836  
; PRIOR FILING DATE: 2005-06-17  
; PRIOR APPLICATION NUMBER: 60/580,969  
; PRIOR FILING DATE: 2004-06-17  
; NUMBER OF SEQ ID NOS: 44  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 11  
; LENGTH: 115  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-11-323-964-11

Query Match 100.0%; Score 597; DB 7; Length 115;  
Best Local Similarity 100.0%; Pred. No. 1.2e-55;  
Matches 115; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MSGEPQTSVAPPPPEVEPGSGVRIIVVEYCEPCGFEATYLELASAVKEQVPGIEISRLG 60  
Db 1 MSGEPQTSVAPPPPEVEPGSGVRIIVVEYCEPCGFEATYLELASAVKEQVPGIEISRLG 60  
Qy 61 GTCAFEIENGQLVFSLKLENGGPPYKDLIEATIRASNGTLEKITNSRPPCVIL 115  
Db 61 GTCAFEIENGQLVFSLKLENGGPPYKDLIEATIRASNGTLEKITNSRPPCVIL 115

##### RESULT 2

US-10-449-902-30557  
; Sequence 30557, Application US/10449902  
; Publication No. US20060123505A1  
; GENERAL INFORMATION:  
; APPLICANT: National Institute of Agrobiological Sciences.



```
; NAME/KEY: unsure
; LOCATION: 13
; OTHER INFORMATION: unknown amino acid
US-10-533-519-233

Query Match      17.0%; Score 101.5; DB 6; Length 87;
Best Local Similarity 32.1%; Pred. No. 0.0017;
Matches 26; Conservative 21; Mismatches 25; Indels 9; Gaps 4;

QY 23 VRIVVEYCEPCGFEATYLELASAVKQYPG-IEI--ESRLGGTGGAPEIEINGQLVFSKLE 79
Db 5 VRVV--YCGAXGYKSKYLQKKLEDFPQRLDICBGTQATGFFVWVAGKLIHSKKK 62

QY 80 NGGFYKED----LIEAIRRA 96
Db 63 GDGYVDTESKFLKVAIAKAA 83

RESULT 6
US-10-953-349-16531
; Sequence 16531, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE OF INVENTION: ENCODED THERBY
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 16531
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (205)..(205)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
US-10-953-349-16531

Query Match      15.6%; Score 93; DB 6; Length 212;
Best Local Similarity 19.1%; Pred. No. 0.039;
Matches 27; Conservative 15; Mismatches 25; Indels 74; Gaps 2;

QY 13 PREE-----VEPGSGVRIVVEYCEPCGFEATYLELASAVKEQYPGIEI----- 55
Db 72 PPEKPISNIATPGLGNTVYINFCSGCSYKGTAVTMKNMLEIALPGTEVILANYPSPK 131

QY 56 -----ESR 58
Db 132 LLSKLVVVQIGVVGVVAGEHIFPMLGFVAPPWPYYNLRANRFGTIASTWLLGNALQSF 191

QY 59 LGGTGAPEIEINGQLVFSKLE 79
Db 192 LQSSGAFEIYFNGLVFSKLE 212

RESULT 7
US-10-953-349-16530
; Sequence 16530, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE OF INVENTION: ENCODED THERBY
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 16530
; LENGTH: 215
```

```
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (228)..(228)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
US-10-953-349-16530

Query Match      15.6%; Score 93; DB 6; Length 235;
Best Local Similarity 19.1%; Pred. No. 0.045;
Matches 27; Conservative 15; Mismatches 25; Indels 74; Gaps 2;

QY 13 PREE-----VEPGSGVRIVVEYCEPCGFEATYLELASAVKEQYPGIEI----- 55
Db 95 PPEKPISNIATPGLGNTVYINFCSGCSYKGTAVTMKNMLEIALPGTEVILANYPSPK 154

QY 56 -----ESR 58
Db 155 LLSKLVVVQIGVVGVVAGEHIFPMLGFVAPPWPYYNLRANRFGTIASTWLLGNALQSF 214

QY 59 LGGTGAPEIEINGQLVFSKLE 79
Db 215 LQSSGAFEIYFNGLVFSKLE 235

RESULT 8
US-10-953-349-16529
; Sequence 16529, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE OF INVENTION: ENCODED THERBY
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 16529
; LENGTH: 236
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (229)..(229)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
US-10-953-349-16529

Query Match      15.6%; Score 93; DB 6; Length 236;
Best Local Similarity 19.1%; Pred. No. 0.045;
Matches 27; Conservative 15; Mismatches 25; Indels 74; Gaps 2;

QY 13 PREE-----VEPGSGVRIVVEYCEPCGFEATYLELASAVKEQYPGIEI----- 55
Db 96 PPEKPISNIATPGLGNTVYINFCSGCSYKGTAVTMKNMLEIALPGTEVILANYPSPK 155

QY 56 -----ESR 58
Db 156 LLSKLVVVQIGVVGVVAGEHIFPMLGFVAPPWPYYNLRANRFGTIASTWLLGNALQSF 215

QY 59 LGGTGAPEIEINGQLVFSKLE 79
Db 216 LQSSGAFEIYFNGLVFSKLE 236

RESULT 9
US-11-395-249-96
; Sequence 96, Application US/11395249
; Publication No. US20060177904A1
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John M.
; APPLICANT: LaVallie, Edward R.
```

```
/ APPLICANT: Collins-Racie, Lisa A.
/ APPLICANT: Evans, Cheryl
/ APPLICANT: Merberg, David
/ APPLICANT: Treacy, Maurice
/ APPLICANT: Agostino, Michael J.
/ APPLICANT: Steiningger II, Robert J.
/ APPLICANT: Spaulding, Vikki
/ APPLICANT: Wong, Gordon G.
/ APPLICANT: Clark, Hilary
/ APPLICANT: Fecthel, Kim
/ TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
/ FILE REFERENCE: 00766.00103.5
/ CURRENT APPLICATION NUMBER: US/11/395,249
/ CURRENT FILING DATE: 2006-04-03
/ NUMBER OF SEQ ID NOS: 240
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 96
/ LENGTH: 137
/ TYPE: prt
/ ORGANISM: Homo sapiens
US-11-395-249-96

Query Match      13.4%; Score 80; DB 7; Length 137;
Best Local Similarity 43.6%; Pred. No. 0.64;
Matches 17; Conservative 9; Mismatches 38; Indels 0; Gaps 0;

QY 55 IESRLGGTGAPRIENGQLVFSKLENGGFPYKDLIEAI 93
Db 81 IENQCMSTGAPRIETLNDVFMVKLESGLHPSMQQLVQIL 119

RESULT 10
US-10-953-349-10235
/ Sequence 10235, Application US/10953349
/ Publication No. US20060107345A1
/ GENERAL INFORMATION:
/ APPLICANT: ALEXANDROV, Nikolai et al.
/ TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
/ FILE REFERENCE: 2750-1579PUS2
/ CURRENT APPLICATION NUMBER: US/10/953,349
/ CURRENT FILING DATE: 2004-09-30
/ NUMBER OF SEQ ID NOS: 40252
/ SOFTWARE: PatentIn version 3.3
/ SEQ ID NO 10235
/ LENGTH: 158
/ TYPE: prt
/ ORGANISM: Arabidopsis thaliana
US-10-953-349-10235

Query Match      13.4%; Score 80; DB 6; Length 158;
Best Local Similarity 30.0%; Pred. No. 0.64;
Matches 24; Conservative 14; Mismatches 38; Indels 4; Gaps 3;

QY 4 EFGQTSVAPPPEVEPGSGVRIVVEYCEPC-GFEATYLELAGAVKQYFG--IEIESRLG 60
Db 54 EAVEEEVEKEEPEVEDPTRTKIVIECHKQCNAFKTRAIQVKEALEGAVPGVTVSLNPEKP 113

QY 61 GTGAPEI-EINGQLVFSKLE 79
Db 114 RRGCFEIREGGQTTFISLLE 133

RESULT 11
US-11-056-355B-106087
/ Sequence 106087, Application US/11056355B
/ Publication No. US20060150283A1
/ GENERAL INFORMATION:
/ APPLICANT: Brover, Vyacheslav
/ APPLICANT: Alexandrov, Nikolai
/ TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
/ FILE REFERENCE: 2750-1590PUS2
/ CURRENT APPLICATION NUMBER: US/11/056,355B
/ CURRENT FILING DATE: 2005-02-14
/ PRIOR APPLICATION NUMBER: 60/544,190
/ PRIOR FILING DATE: 2004-02-13
/ NUMBER OF SEQ ID NOS: 119966
/ SEQ ID NO 117326
/ LENGTH: 158
/ TYPE: prt
/ ORGANISM: Arabidopsis thaliana
US-11-056-355B-117326

Query Match      13.4%; Score 80; DB 7; Length 158;
Best Local Similarity 30.0%; Pred. No. 0.64;
Matches 24; Conservative 14; Mismatches 38; Indels 4; Gaps 3;

QY 4 EFGQTSVAPPPEVEPGSGVRIVVEYCEPC-GFEATYLELAGAVKQYFG--IEIESRLG 60
Db 54 EAVEEEVEKEEPEVEDPTRTKIVIECHKQCNAFKTRAIQVKEALEGAVPGVTVSLNPEKP 113

QY 61 GTGAPEI-EINGQLVFSKLE 79
Db 114 RRGCFEIREGGQTTFISLLE 133

RESULT 12
US-11-056-355B-117326
/ Sequence 117326, Application US/11056355B
/ Publication No. US20060150283A1
/ GENERAL INFORMATION:
/ APPLICANT: Brover, Vyacheslav
/ APPLICANT: Alexandrov, Nikolai
/ TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
/ FILE REFERENCE: 2750-1590PUS2
/ CURRENT APPLICATION NUMBER: US/11/056,355B
/ CURRENT FILING DATE: 2005-02-14
/ PRIOR APPLICATION NUMBER: 60/544,190
/ PRIOR FILING DATE: 2004-02-13
/ NUMBER OF SEQ ID NOS: 119966
/ SEQ ID NO 117326
/ LENGTH: 158
/ TYPE: prt
/ ORGANISM: Arabidopsis thaliana
US-11-056-355B-117326

Query Match      13.4%; Score 80; DB 7; Length 158;
Best Local Similarity 30.0%; Pred. No. 0.64;
Matches 24; Conservative 14; Mismatches 38; Indels 4; Gaps 3;

QY 4 EFGQTSVAPPPEVEPGSGVRIVVEYCEPC-GFEATYLELAGAVKQYFG--IEIESRLG 60
Db 54 EAVEEEVEKEEPEVEDPTRTKIVIECHKQCNAFKTRAIQVKEALEGAVPGVTVSLNPEKP 113

QY 61 GTGAPEI-EINGQLVFSKLE 79
Db 114 RRGCFEIREGGQTTFISLLE 133

RESULT 13
US-10-953-349-15768
/ Sequence 15768, Application US/10953349
/ Publication No. US20060107345A1
/ GENERAL INFORMATION:
/ APPLICANT: ALEXANDROV, Nikolai et al.
/ TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
/ FILE REFERENCE: 2750-1590PUS2
```

```
/ CURRENT APPLICATION NUMBER: US/11/056,355B
/ CURRENT FILING DATE: 2005-02-14
/ PRIOR APPLICATION NUMBER: 60/544,190
/ PRIOR FILING DATE: 2004-02-13
/ NUMBER OF SEQ ID NOS: 119966
/ SEQ ID NO 106087
/ LENGTH: 158
/ TYPE: prt
/ ORGANISM: Arabidopsis thaliana
/ FEATURE:
/ NAME/KEY: peptide
/ LOCATION: (1)..(158)
/ OTHER INFORMATION: Ceres Seq. ID no. 13618700
US-11-056-355B-106087

Query Match      13.4%; Score 80; DB 7; Length 158;
Best Local Similarity 30.0%; Pred. No. 0.64;
Matches 24; Conservative 14; Mismatches 38; Indels 4; Gaps 3;

QY 4 EFGQTSVAPPPEVEPGSGVRIVVEYCEPC-GFEATYLELAGAVKQYFG--IEIESRLG 60
Db 54 EAVEEEVEKEEPEVEDPTRTKIVIECHKQCNAFKTRAIQVKEALEGAVPGVTVSLNPEKP 113

QY 61 GTGAPEI-EINGQLVFSKLE 79
Db 114 RRGCFEIREGGQTTFISLLE 133
```

```
RESULT 12
US-11-056-355B-117326
/ Sequence 117326, Application US/11056355B
/ Publication No. US20060150283A1
/ GENERAL INFORMATION:
/ APPLICANT: Brover, Vyacheslav
/ APPLICANT: Alexandrov, Nikolai
/ TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
/ FILE REFERENCE: 2750-1590PUS2
/ CURRENT APPLICATION NUMBER: US/11/056,355B
/ CURRENT FILING DATE: 2005-02-14
/ PRIOR APPLICATION NUMBER: 60/544,190
/ PRIOR FILING DATE: 2004-02-13
/ NUMBER OF SEQ ID NOS: 119966
/ SEQ ID NO 117326
/ LENGTH: 158
/ TYPE: prt
/ ORGANISM: Arabidopsis thaliana
/ FEATURE:
/ NAME/KEY: peptide
/ LOCATION: (1)..(158)
/ OTHER INFORMATION: Ceres Seq. ID no. 13618700
US-11-056-355B-117326
```

```
Query Match      13.4%; Score 80; DB 7; Length 158;
Best Local Similarity 30.0%; Pred. No. 0.64;
Matches 24; Conservative 14; Mismatches 38; Indels 4; Gaps 3;

QY 4 EFGQTSVAPPPEVEPGSGVRIVVEYCEPC-GFEATYLELAGAVKQYFG--IEIESRLG 60
Db 54 EAVEEEVEKEEPEVEDPTRTKIVIECHKQCNAFKTRAIQVKEALEGAVPGVTVSLNPEKP 113

QY 61 GTGAPEI-EINGQLVFSKLE 79
Db 114 RRGCFEIREGGQTTFISLLE 133
```

```
RESULT 13
US-10-953-349-15768
/ Sequence 15768, Application US/10953349
/ Publication No. US20060107345A1
/ GENERAL INFORMATION:
/ APPLICANT: ALEXANDROV, Nikolai et al.
/ TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
/ FILE REFERENCE: 2750-1590PUS2
```

```
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 15767
; LENGTH: 184
; TYPE: PRT
; ORGANISM: Glycine max
US-10-953-349-15767

Query Match      13.4%; Score 80; DB 6; Length 184;
Best Local Similarity 28.3%; Pred. No. 0.78;
Matches 26; Conservative 12; Mismatches 32; Indels 22; Gaps 5

Qy    4   EPQG--TSVAPPEEVEP-----GSGVRIVVEYCEPCGFATYLELASAVKEQYPGIEIES 57
       :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    44  QPSQPATPPVPFPAAEPKQPQVSTSPSPVKTFSDVLAFSGPAPE-----RING 91
       :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy    58  RL---GGTGAFEI-I-NGQLVFSLKLENGGPY 85
       :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    92  RLANIGFVGMAVEVAKGQGVFEQISNGGI PW 123
       :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Search completed: December 6, 2006, 10:51:27
Job time : 75 secs
```

```

; TITLE OF INVENTION: ENCODED THERY
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 15768
; LENGTH: 176
; TYPE: PRT
; ORGANISM: Glycine max
US-10-953-349-15768

Query Match      13.4%; Score 80; DB 6; Length 176;
Best Local Similarity 28.3%; Pred. No. 0.74;
Matches 26; Conservative 12; Mismatches 32; Indels 22; Gaps 5;

Qy      4  EPG- -TSVAPPPPEVEP- - - -GSGVRIVRYCBPCGFEATYLELAGSAVKQYPCGIEIES 57
      36  QPSQPATPTVTPPPAEPKQPVSTSPKVVTKFSDVLAFGPAPE- - - - - - - - - - -RING 83
      58  RL- - - -GTTGAFETEI -NGQLVFSKLENGGPPY 85
      84  RLAMIGFVGMAVEAVKQGVFEQISNGGIPW 115

```

RESULT 14

US-11-056-355B-53080

US-11-056-355B-53080

Sequence 53080, Application US/11056355B

Publication No. US20060150283A1

GENERAL INFORMATION:

APPLICANT: Brover, Vyacheslav

APPLICANT: Alexandrov, Nikolai

TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding

TITLE OF INVENTION: Polypeptides Encoded Thereby

FILE REFERENCE: 2750-1590PUS2

CURRENT APPLICATION NUMBER: US/11/056,355B

CURRENT FILING DATE: 2005-02-14

PRIOR APPLICATION NUMBER: 60/544,190

PRIOR FILING DATE: 2004-02-13

NUMBER OF SEQ ID NOS: 119966

SEQ ID NO 53080

LENGTH: 176

TYPE: prt

ORGANISM: Glycine max

FEATURE:

NAME/KEY: peptide

LOCATION: (1)..(176)

OTHER INFORMATION: Ceres Seq. ID no. 13637661

US-11-056-355B-53080

```

Query Match      13.4%; Score 80; DB 7; Length 176;
Best Local Similarity 28.3%; Pred. No. 0.74;
Matches 26; Conservative 12; Mismatches 32; Indels 22; Gaps 5;

Qy   4 EPQG--TSVAPPPPEVP-----GSGRVIVVCPCGCFEATYLELASAVKEQYPCIEIES 57
Db       :| |:| |:| |:| ::::| |:| |:| |:| |
Qy   36 QPSQPATFVTTPPAEPKPPQPVSTSPKVKTFSDVLAFSGPAPE-----RING 83
Db       :| |:| |:| |:| ::::| |:| |:| |:| |
Qy   58 RL----GGTGAFEIET-NGQLVFSSKLENGGFYP 85
Db       :| |:| |:| |:| ::::| |:| |:| |:| |
Qy   84 RLMIGFGMAVEAVAKGGGVFEQISNGGIWF 115
Db       :| |:| |:| |:| ::::| |:| |:| |:| |

RESULT 15
US-10-953-349-15767
; Sequence 15767, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953.349
```

was a blank (1870)

GenCore version 5.1.1.9  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: December 6, 2006, 10:37:27 ; Search time 23 Seconds  
(without alignments)  
481.084 Million cell updates/sec

Title: US-09-824-787B-2

Perfect score: 597

Sequence: 1 MSGEPGQTSVAPPPEVEPG.....ASNGETLEKITSRPPCVIL 115

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 80: \*  
1: pir1: \*  
2: pir2: \*  
3: pir3: \*  
4: pir4: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	92	15.4	255	2 T19747	hypothetical prote
2	92	15.4	258	2 T19755	hypothetical prote
3	90	15.1	101	2 AD2604	conserved hypothet
4	90	15.1	101	2 D97386	hypothetical prote
5	83	13.9	96	2 D83228	hypothetical prote
6	81	13.6	97	2 C82257	selenoprotein W-re
7	80	13.4	243	2 A72669	probable glutaredo
8	79	13.2	395	2 E86152	chalcone synthase
9	79	13.2	469	2 D72722	probable MRSA prot
10	78.5	13.1	101	2 AD3542	hypothetical prote
11	78	13.1	3670	2 T36249	CDA peptide synthet
12	76	12.7	232	2 T21526	hypothetical prote
13	75.5	12.6	393	1 E64485	phosphoribosylglyc
14	74	12.4	698	2 AD2985	oxidoreductase Atu
15	74	12.4	729	2 C98298	probable oxidoredu
16	74	12.4	969	2 A75634	McrB-related prote
17	73	12.2	163	2 D87713	protein-export pro
18	72.5	12.1	209	2 H83799	two-component resp
19	72.5	12.1	428	2 B64301	2,3-bisphosphoglyc
20	72	12.1	348	2 T18230	alcohol dehydrogen
21	72	12.1	525	2 F70068	gamma-glutamyltran
22	72	12.1	938	2 A13417	glutamate-ammonia
23	72	12.1	987	2 I48953	eph-related recept
24	72	12.1	1507	2 D97106	large chain of NAD
25	71.5	12.0	347	2 G83171	hypothetical prote
26	71	11.9	413	2 T03270	probable histidino
27	71	11.9	448	1 C69594	adenosylmethionine
28	70.5	11.8	282	2 F86577	elongation factor
29	70.5	11.8	467	2 C72601	probable tryptopha

ALIGNMENTS

RESULT 1

T19747

hypothetical protein C35C5.3a - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 18-Feb-2000

C:Accession: T19747

R:White, S.

submitted to the EMBL Data Library, August 1996

A:Reference number: Z19173

A:Accession: T19747

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-255 <WIL>

A:Cross-references: UNIPARC:UPI0000164317; EMBL:Z78417; PIDN: CAB01684.1; GSPDB:GN000028;

A:Experimental source: clone C35C5

C:Genetics:

A:Gene: CESP:C35C5.3a

A:Map position: X

A:Introns: 11/3; 34/1; 64/2; 100/2; 141/3

Query Match 15.4%; Score 92; DB 2; Length 255;

Best Local Similarity 18.0%; Pred. No. 0.2;

Matches 24; Conservative 18; Mismatches 23; Indels 68; Gaps 2;

QY 29 YCEPCGFEATYLELASAVKEQYEGIEISR-----LGG- 61

DB 99 YCVSCGYKQAFDQFTTFAKEKYPNMFIEGANFAPVLWKAYVAQALSFVKNAVLLVLGGI 158

QY 62 -----TGAFRIEINGQLVFSKLEN 80

DB 159 NPFERFGLYPQILQHAHGNKMSSCMLVFMLGNLVFQSLISTGAFEVILGNEQIWSKIES 218

QY 81 GGFPVEKDLIEAI 93

DB 219 GRVPSQPQFMQLI 231

RESULT 2

T19755

hypothetical protein C35C5.3b - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 18-Feb-2000

C:Accession: T19755

R:White, S.

submitted to the EMBL Data Library, August 1996

A:Reference number: Z19173

A:Accession: T19755

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-258 <WIL>

A:Cross-references: UNIPARC:UPI0000164318; EMBL:Z78417; PIDN: CAB01692.1; GSPDB:GN000028;

A:Experimental source: clone C35C5

chalcone isomerase  
hypothetical prote  
probable 3-hydroxy  
outer membrane pro  
protein F9C16.9 [i  
hypothetical prote  
involved in spore  
molybdenum cofacto  
phosphoribosylglyc  
FMN oxidoreductase  
proline dehydrogen  
hypothetical prote  
hypothetical prote  
molybdopterin conv  
methylated-DNA- [pr  
hypothetical prote

```
C;Genetics:
A;Gene: CESP:C35CS.3b
A;Map position: X
A;Introns: 11/3; 34/1; 67/2; 103/2; 144/3

Query Match      15.4%; Score 92; DB 2; Length 258;
Best Local Similarity 18.0%; Pred. No. 0.21;
Matches 24; Conservative 18; Mismatches 23; Indels 68; Gaps 2;

Qy      29 YCEPGCFEATYLELAGAVKEQYPCIEIESR-----LGG- 61
       |||::|||::|||::|||::|||::|||::|||::|||
Db      102 YCVSCGYKQAFDQFTTFAKEKYENMPLEGANFAPVLWKAYVAQALSFVMMAVLVULGGI 161
Qy      62 -----TGAFEIENGQLVFESKLEN 80
       |||::|||::|||::|||::|||::|||::|||::|||
Db      162 NPFERFGLYPQILQHAGHGMSSCMLVFMGLNLVEQSLLSTGAFEVYLGNQIWSKIES 221
Qy      81 GGFPEKDLIEAI 93
Db      222 GRVPSPQEFMQLI 234

RESULT 3
AD2604
conserved hypothetical protein Atu0228 [imported] - Agrobacterium tumefaciens (strain C58)
C;Species: Agrobacterium tumefaciens
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C;Accession: AD2604
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.;
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClellan,
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A;Reference number: AB2577; PMID:21608550; PMID:11743193
A;Accession: AD2604
A>Status: preliminary
A:Molecule type: DNA
A;Residues: 1-101<KUR>
A;Cross-references: UNIPROT:QBUIR5; UNIPARC:UPI000000D173B; GB:AE008698; PIDN:AAL41250.1;
A;Experimental source: strain C58 (Dupont)
C;Genetics:
A;Gene: Atu0228
A;Map position: circular chromosome

Query Match      15.1%; Score 90; DB 2; Length 101;
Best Local Similarity 29.6%; Pred. No. 0.11;
Matches 24; Conservative 17; Mismatches 34; Indels 6; Gaps 3;

Qy      16 EVEPGSVRIIVVEYCEPGCFEATYLELAGAVKEQYPC-IEIESRLGGTGA-FEISINGOL 73
       |||::|||::|||::|||::|||::|||::|||::|||
Db      3 ETKP----RIAIRCYCQNWLNRAGWMAQEILQTIFASDIGEVSLPSTGGFLFEITVDGTI 58
Qy      74 VFSKLENGGFPYEKDIEAIR 94
       ::::|::|::|::|::|::|::|::|::|::|
Db      59 IWERKRDGGFPGPKEIKQRIR 79

RESULT 4
D97386
hypothetical protein AGR_C_387 [imported] - Agrobacterium tumefaciens (strain C58, Cerec
C;Species: Agrobacterium tumefaciens
C;Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
C;Accession: D97386
R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorollo, B.; Goldman,
A.; Liu, F.; Wellam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Leppas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A;Reference number: A97359; PMID:21608551; PMID:11743194
A;Accession: D97386
A>Status: preliminary
A:Molecule type: DNA
```

A;Residues: 1-101 <KUR>  
A;Cross-references: UNIPROT:Q8UIR5; UNIPARC:UPI00000D173B; GB:AE007869; PIDN:AAK86045.1.1;  
C;Genetics:  
A;Gene: AGR\_C.387  
A;Map position: circular chromosome

Query Match 15.1%; Score 90; DB 2; Length 101;  
Best Local Similarity 25.6%; Pred. No. 0.11;  
Matches 24; Conservative 17; Mismatches 34; Indels 6; Gaps 3;

Qy 16 EVEPGSGVRIWVECEPCGFATYLELASAVKEQYPG-IEIESRLGGTGA-PEIEINGQL 73  
Db 3 ETKP-----RIAIRYCTQCNWLLRAGWMAQEILQTFASDIGEVSLIPSTGGLFEITVDGTI 58  
Qy 74 VFSLKLENGFPFYEKDLIEAIR 94  
Db 59 IWERKRDGGFPGPKEKQRIIR 79

RESULT 5  
D83228  
hypothetical protein PA3338 [imported] - Pseudomonas aeruginosa (strain PAO1)  
C;Species: Pseudomonas aeruginosa  
C;Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 09-Jul-2004  
C;Accession: D83228  
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Bri-  
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,  
.; Lory, S.; Olson, M.V.  
Nature 406, 959-964, 2000  
A;Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic patho-  
A;Reference number: A82950; MUID:20437337; PMID:10984043  
A;Accession: D83228  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-96 <STO>  
A;Cross-references: UNIPROT:Q9HYQ7; UNIPARC:UPI00000C59C0; GB:AE004756; GB:AE004091; NID:  
A;Experimental source: strain PAO1  
C;Genetics:  
A;Gene: PA3338

Query Match 13.9%; Score 83; DB 2; Length 96;  
Best Local Similarity 31.2%; Pred. No. 0.49;  
Matches 25; Conservative 13; Mismatches 36; Indels 6; Gaps 3;

Qy 19 PGSGVRIWVECEPCGPF--EATYL--ELASAVKEQYPGIEIESRLGGTGAPIEINGQLV 74  
Db 2 PTAKEPIVITYCTCQWLLRAWLAQELLSFADDLGKVCLEPGTG--GVFRIITCDGVQV 59  
Qy 75 FSKLENGFPFYEKDLIEAIR 94  
Db 60 WERKADGGFPPEAKALKQQRV 79

RESULT 6  
C82257  
selenoprotein W-related protein VC0982 [imported] - Vibrio cholerae (strain N16961 serog:  
C;Species: Vibrio cholerae  
C;Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 09-Jul-2004  
C;Accession: C82257  
R;Haideberg, J.F.; Eissen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.; J-  
charadon, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P.;  
1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.  
Nature 406, 477-483, 2000  
A;Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.  
A;Reference number: A82035; MUID:20406833; PMID:10952301  
A;Accession: C82257  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-97 <HEI>  
A;Cross-references: UNIPROT:Q9KTC1; UNIPARC:UPI00000C2E43; GB:AE004179; GB:AE003852; NID:  
A;Experimental source: serogroup O1; strain N16961; biotype El Tor  
C;Genetics:  
A;Gene: VC0982







R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.  
erage, G.; Gallet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClell  
; Karp, P.; Romero, P.; Zhang, S.  
Science 294, 2317-2323, 2001  
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,  
ster, E.W.  
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.  
A:Reference number: AB2577; MUID:21608550; PMID:11743193  
A:Accession: AD2985  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-698 <KUR>  
A:Cross-references: UNIPROT:Q8UA89; UNIPARC:UPI000016480F; GB:AE008689; PIDN:AAL44298.1;  
A:Experimental source: strain C58 (Dupont)  
C:Genetics:  
A:Gene: Atu3485  
A:Map position: linear chromosome

Query Match 12.4%; Score 74; DB 2; Length 698;  
Best Local Similarity 22.9%; Pred. No. 35;  
Matches 25; Conservative 20; Mismatches 38; Indels 26; Gaps 3;  
QY 16 EVEPGSGVRIVVEYCEPC-----GFEATYLELASAVKEQYPGIEIESRLGG----- 61  
Db 53 EIDPLTGEKVRVMWVGSGDVGTIKLDGFATLYLEKLEALKGIYKGVDEDEDMVGFLPH 112  
QY 62 -----TGAPEIEINGQLVFSKLENGGFPPEYKDLIEAIRRASNGETL 102  
Db 113 CTNLPRAASIDTPLHGFVDFVDH-----MHPDAIIAIAASKNSKEL 157

## RESULT 15

C98298  
probable oxidoreductase [imported] - Agrobacterium tumefaciens (strain C58, Cereon)  
C:Species: Agrobacterium tumefaciens  
C:Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 09-Jul-2004  
C:Accession: C98298  
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorollo, B.; Goldman,  
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;  
Science 294, 2323-2328, 2001  
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum  
A:Reference number: A97359; MUID:21608551; PMID:11743194  
A:Accession: C98298  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-729 <KUR>  
A:Cross-references: UNIPROT:Q8UA89; UNIPARC:UPI00000D2209; GB:AE007870; PIDN:AAK89909.1;  
C:Genetics:  
A:Gene: AGR\_L\_2690  
A:Map position: linear chromosome

Query Match 12.4%; Score 74; DB 2; Length 729;  
Best Local Similarity 22.9%; Pred. No. 37;  
Matches 25; Conservative 20; Mismatches 38; Indels 26; Gaps 3;  
QY 16 EVEPGSGVRIVVEYCEPC-----GFEATYLELASAVKEQYPGIEIESRLGG----- 61  
Db 84 EIDPLTGEKVRVMWVGSGDVGTIKLDGFATLYLEKLEALKGIYKGVDEDEDMVGFLPH 143  
QY 62 -----TGAPEIEINGQLVFSKLENGGFPPEYKDLIEAIRRASNGETL 102  
Db 144 CTNLPRAASIDTPLHGFVDFVDH-----MHPDAIIAIAASKNSKEL 188

Search completed: December 6, 2006, 10:40:16  
Job time : 25 secs

**This Page Blank (uspic)**

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: December 6, 2006, 10:26:47 ; Search time 148 Seconds  
(without alignments)  
718.762 Million cell updates/sec

Title: US-09-824-787B-2

Perfect score: 597

Sequence: 1 MSGEPGQTSVAPPPBEVPG.....ASNGETLEKITSRPPCVIL 115

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2849598 seqs, 92501592 residues

Total number of hits satisfying chosen parameters: 2849598

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt 7.2.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	597	100.0	115	2	Q9BRT3_HUMAN	Q9Brt3 homo sapien
2	537	89.9	115	2	Q9C086_MOUSE	Q9c086 m adult mal
3	435.5	72.9	126	2	Q5ZIH7_CHICK	Q5zih7 gallus gall
4	288	48.2	95	2	Q3ZLC7_OREMO	Q3zic7 oreochromis
5	273	45.7	95	2	Q802F5_BRARE	Q802f5 brachydanio
6	221.5	37.1	94	2	Q802G8_BRARE	Q802g8 brachydanio
7	218	36.5	89	2	Q4T504_TETNG	Q4t504 tetraodon n
8	148.5	24.9	95	2	Q9VRA0_DROME	Q9vra0 drosophila
9	137.5	23.0	88	2	Q8H6T4_CHLRE	Q8h6t4 chlamydomon
10	137	22.9	80	2	Q4JBM6_SULAC	Q4jbm6 sulfolobus
11	132.5	22.2	93	2	Q7QFL8_ANOGA	Q7qfl8 anopheles g
12	119	19.9	232	2	Q8S227_ORYSA	Q8s227 oryza sativ
13	116.5	19.5	104	2	Q3QTH4_RHOBO	Q3qth4 silicibacte
14	114.5	19.2	244	1	HSPE_HETGL	Q9bn19 heterodera
15	112.5	18.8	86	1	SEPW1_HUMAN	P63302 homo sapien
16	112.5	18.8	86	1	SEPW1_MACMU	P63303 macaca mula
17	112.5	18.8	86	1	SEPW1_PIG	Q95kl4 sus scrofa
18	112.5	18.8	86	2	Q802G9_BRARE	Q802g9 brachydanio
19	112.5	18.8	87	2	Q5NVB2_PONPY	Q5nvb2 pongo pygma
20	111.5	18.7	86	2	Q568W0_BRARE	Q568w0 brachydanio
21	110.5	18.5	209	1	SELT_ARATH	Q9stz2 arabidopsis
22	110.5	18.5	346	1	SELV_HUMAN	P59797 homo sapien
23	109	18.3	328	2	Q5FWB9_MOUSE	Q5fwb9 mus musculus
24	107.5	18.0	87	1	SEPW1_MOUSE	P63300 mus musculus
25	107.5	18.0	87	1	SEPW1_RAT	P63301 rattus norv
26	104	17.4	257	2	Q615I7_ORYSA	Q615i7 oryza sativ
27	103.5	17.3	228	2	Q8W1E5_ARATH	Q8wie5 arabidopsis
28	101	16.9	95	2	Q3XA62_METFL	Q3xa62 methylobaci
29	100	16.8	97	2	Q2KDL9_RHIEI	Q2kdl9 rhizobium e
30	99	16.6	92	2	Q8DFL8_VIBVU	Q8dfl8 vibrio vuln
31	99	16.6	94	2	Q7MMS1_VIBVY	Q7mms1 vibrio vuln

32	98.5	16.5	86	1	SEPW1_SHEEP	O19097 ovis aries
33	98	16.4	92	2	Q36PV6_MARHY	Q36pv6 marinobacte
34	95	15.9	94	2	Q3WKN7_9RHIZ	Q3wkn7 mesorhizobi
35	94	15.7	92	2	Q87RH7_VIBPA	Q87rh7 vibrio para
36	94	15.7	103	2	Q60CA7_METCA	Q60ca7 methylococc
37	93	15.6	93	2	Q92SU7_RHIME	Q92su7 rhizobium m
38	93	15.6	201	2	Q7PVR5_ANOGA	Q7pvr5 anopheles g
39	92.5	15.5	163	2	Q802F2_BRARE	Q802f2 brachydanio
40	92	15.4	81	2	Q7QZL6_GIALA	Q7qz16 giardia lam
41	92	15.4	247	1	SELT1_CAREL	Q9u3n5 caenorhabdi
42	92	15.4	247	2	Q610U0_CABER	Q610u0 caenorhabdi
43	90	15.1	90	2	Q5YEV1_CHUS6	Q5yev1 chlorarachn
44	90	15.1	101	2	Q8UIR5_AGR56	Q8uir5 agrobacteri
45	89.5	15.0	1125	2	Q5U4C0_MOUSE	Q5u4c0 mus musculu

## ALIGNMENTS

RESULT 1  
Q9BRT3\_HUMAN PRELIMINARY; PRT; 115 AA.  
AC Q9BRT3;  
DT 01-JUN-2001, integrated into UniProtKB/TrEMBL.  
DT 01-JUN-2001, sequence version 1.  
DT 07-FEB-2006, entry version 21.  
DE Chromosome 17 open reading frame 37 (XTP4) (C35 protein).  
GN Name=C17orf37; Synonyms=XTP4;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;  
OC Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=Brain;  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Dege J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalish D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [2]  
RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=Brain;  
RA Director MGC Project;  
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP NUCLEOTIDE SEQUENCE.  
RA Liu Y., Cheng J., Lu Y., Wang G., Zhang L., Chen J., Li L.;  
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP NUCLEOTIDE SEQUENCE.  
RA Evans E.E., Henn A.D., Luhowskyj S., Paris M.J., Borrello M.A.,  
RA Smith E.S., Sahasrabudhe D.M., Zauderer M.;  
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.  
CC -----  
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>

```

CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; BC006006; AAH06006.1; -; mRNA.
DR EMBL; AY508253; AAO85461.1; -; mRNA.
DR EMBL; AY508814; AAO92035.1; -; mRNA.
DR Ensembl; ENSG00000141741; Homo sapiens.
DR HGNC; HGNC:28230; C17orf17.
DR LinkHub; Q9BR13; -.
DR GO; GO:0008430; F:selenium binding; IEA.
DR GO; GO:0045454; P:cell redox homeostasis; IEA.
DR InterPro; IPR011893; CXXU selWTH.
DR TIGRFAMs; TIGR021174; CXXU selWTH; 1.
DR SEQUENCE 115 AA; 12403 MW; 5D8B911C0P23DDC1 CRC64;
QY Query Match 100.0%; Score 597; DB 2; Length 115;
Best Local Similarity 100.0%; Pred. NO. 8e-51;
Matches 115; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSGEQQTQSVAPPPVEPGSGVRIIVVEYCEPGFEATYLEASAVKEQYPGIEIESRLG 60
DB 1 MSGEQQTQSVAPPPVEPGSGVRIIVVEYCEPGFEATYLEASAVKEQYPGIEIESRLG 60
QY 61 GTGAPFIEINGQLVFSKLENGGPPYKDLIEAIRRASNGETLEKITSNPPPCVIL 115
DB 61 GTGAPFIEINGQLVFSKLENGGPPYKDLIEAIRRASNGETLEKITSNPPPCVIL 115

RESULT 2
Q9C086 MOUSE PRELIMINARY; PRT; 115 AA.
AC Q9C086;
DT 01-JUN-2001, integrated into UniProtKB/TrEMBL.
DT 01-JUN-2001, sequence version 1.
DT 07-FEB-2006, entry version 24.
DE Adult male tongue cDNA, RIKEN full-length enriched library,
DE clone:2310051814 product:hypothetical protein, full insert sequence
DE (Adult male brain cDNA, RIKEN full-length enriched library,
DE clone:0710001816 product:hypothetical protein, full insert sequence)
DE (10 day old male pancreas cDNA, RIKEN full-length enriched library,
DE clone:1810046319 product:hypothetical protein, full insert sequence)
DE (Adult male aorta and vein cDNA, RIKEN full-length enriched library,
DE clone:A530099234 product:hypothetical protein, full insert sequence)
DE (RIKEN cDNA 1810046319).
GN Name=1810046319R1K;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Aorta and vein, Brain, Pancreas, and Tongue;
RX PubMed=16141072; DOI=10.1126/science.1112014;
RA Carninci P., Katayama T., Katayama S., Gough J., Frith M.C., Maeda N.,
RA Oyama R., Ravasi T., Lenhard B., Wells C., Kodzius R., Shimokawa K.,
RA Bajic V.B., Brenner S.E., Batalov S., Forrest A.R., Zavolan M.,
RA Davis M.J., Wilmong L.G., Aldinis V., Allen J.E.,
RA Ambesi-Impombato A., Apweiler K.W., Aturaliya R.N., Bailey T.L.,
RA Bansal M., Baxter L., Beisel K.W., Bersano T., Bono H., Chalk A.M.,
RA Chiu K.P., Choudhary V., Christoffels A., Clutterbuck D.R.,
RA Crowe M.L., Dalla E., Dalrymple B.P., de Bono B., Della Gatta G.,
RA Di Bernar C.F., Fukushima T., Furuno M., Futaki S., Gariboldi M.,
RA Georgii-Hemming P., Gingeras T.R., Gojobori T., Green R.E.,
RA Gustincich S., Harbers M., Hayashi Y., Hensch T.K., Hirokawa N.,
RA Hill D., Hummel L., Iacono M., Iakeo K., Iwano A., Ishikawa T.,

```

```

RA Jakt M., Kanapin A., Katoh M., Kawasawa Y., Kelson J., Kitamura H.,
RA Kitano H., Kollias G., Krishnan S.P., Kruger A., Kummerfeld S.K.,
RA Kurochkin I.V., Lareau L.F., Lazarevic D., Lipovich L., Liu J.,
RA Liuni S., McWilliam S., Madan Babu M., Madera M., Marchionni L.,
RA Mottagui-Tabar S., Mulder N., Nakano N., Nakachi H., Ng P.,
RA Nilsson R., Nishiguchi S., Nishikawa S., Nori F., Ohara O.,
RA Okazaki Y., Orlando V., Pang K.C., Pavan W.J., Pavese G., Pesole G.,
RA Petrovsky N., Piazza S., Reed J., Reid J.F., Ring B.Z., Ringwald M.,
RA Rost B., Ruan Y., Salzberg S.L., Sanderlin A., Schneider C.,
RA Schonbach C., Sekiguchi K., Semple C.A., Sene S., Sessa L., Sheng Y.,
RA Shibata Y., Shimada H., Shimada K., Silva D., Sinclair B.,
RA Sperling S., Stupka E., Sugura K., Sultana R., Takenaka Y., Taki K.,
RA Tammoja K., Tan S.L., Tang S., Taylor M.S., Tegner J., Teichmann S.A.,
RA Ueda H.R., van Nimwegen E., Verardo R., Wei C.L., Yagi K.,
RA Yamanishi H., Zabarovsky E., Zhu S., Zimmer A., Hide W., Bult C.,
RA Grimmond S.M., Teasdale R.D., Liu E.T., Brusic V., Quackenbush J.,
RA Wahlestedt C., Mattick J.S., Hume D.A., Kai C., Sasaki D., Tomaru Y.,
RA Fukuda S., Kanamori-Katayama M., Suzuki M., Aoki J., Arakawa T.,
RA Lida J., Inamura K., Itoh M., Kato T., Kawaji H., Kawagashira N.,
RA Kawashima T., Kojima M., Kondo S., Konno H., Nakano K., Ninomiya N.,
RA Nishio T., Okada M., Plesey C., Shibata K., Shiraki T., Suzuki S.,
RA Tagami M., Waki K., Watahiki A., Okamura-Oho Y., Suzuki H., Kawai J.,
RA Hayashizaki Y.;
RT "The transcriptional landscape of the mammalian genome.";
RL Science 309:1559-1563 (2005).
RN [3]
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Aorta and vein, Brain, Pancreas, and Tongue;
RX PubMed=16141073; DOI=10.1126/science.1112009;
RG RIKEN Genome Exploration Research Group, and Genome Science Group
RG (Genome Network Core Team) and the FANTOM Consortium;
RT "Antisense Transcription in the Mammalian Transcriptome.";
RL Science 309:1564-1566 (2005).
RN [4]
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Aorta and vein, Brain, Pancreas, and Tongue;
RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA Nikaide I., Oatso N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA Balderelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Bradt D., Brusci V., Chothia C., Corbani L.E., Cousins S.,
RA Dalla E., Dragni T.A., Fletcher C., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Kongsawa A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sanderlin A., Schneider C., Semple C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA Hironaka-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Iehi Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shingawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of
RL Nature 420:563-573 (2002).
RN [5]
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Aorta and vein, Brain, Pancreas, and Tongue;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,

```

RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,  
RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,  
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavani T.,  
RA Fleischnann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,  
RA Schriml L.M., Stauble F., Suzuki R., Tomita M., Wagner L., Washio T.,  
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.P.,  
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
RA Wynehaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohteuki S.,  
RA Hayashizaki Y.,  
RT "Functional annotation of a full-length mouse cDNA collection.";  
RL Nature 409:685-690(2001).  
RN [6]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=C57BL/6J; TISSUE=Aorta and vein, Brain, Pancreas, and Tongue;  
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;  
RA Carninci P., Shibata Y., Hayatsu M., Sugahara Y., Shibata K., Itoh M.,  
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.,  
RT "Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new genes.";  
RL Genome Res. 10:1617-1630(2000).  
RN [7]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=C57BL/6J; TISSUE=Aorta and vein, Brain, Pancreas, and Tongue;  
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;  
RA Shibata K., Itoh M., Aizawa K., Nagao S., Sasaki N., Carninci P.,  
RA Konno H., Akiyama J., Nishi K., Kusunagi T., Tashiro H., Itoh M.,  
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,  
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,  
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,  
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsumura S., Kawai J.,  
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.,  
RT "RIKEN integrated sequence analysis (RISA) system-384-format  
sequencing pipeline with 384 multicapillary sequencer.";  
RL Genome Res. 10:11757-11771(2000).  
RN [8]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=C57BL/6J; TISSUE=Brain, Pancreas, and Tongue;  
RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,  
RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,  
RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,  
RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,  
RA Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,  
RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,  
RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,  
RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,  
RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,  
RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,  
RA Muramatsu M., Hayashizaki Y.,  
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
RN [9]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=C57BL/6J; TISSUE=Aorta and vein;  
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,  
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,  
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,  
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,  
RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,  
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,  
RA Nishi K., Nomura K., Numazaki R., Ohno M., Oheato N., Okazaki Y.,  
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,  
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,  
RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,  
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.,  
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
RN [10]  
RP NUCLEOTIDE SEQUENCE.

RC STRAIN=FVB/N; TISSUE=Liver;  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Krausner R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D.,  
RA Klein R.D., Collins F.S., Wagner L., Shemmen C.M., Schaefer C.F., Bhat N.K.,  
RA Altschul S.F., Zeeb B., Buetow K.H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Hopkins R.F., Jordan H., Marusina K., Farmer A., Rubin G.M., Hong L.,  
RA Diatchenko L., Marusina K., Farmer A., Rubin G.M., Hong L.,  
RA Diatchenko L., Marusina K., Farmer A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
Query Match 89.9%; Score 537; DB 2; Length 115;  
Best Local Similarity 89.6%; Pred. NO. 6.7e-45;  
Matches 103; Conservative 4; Mismatches 8; Indels 0; Gaps 0;  
QY 1 MSGEPCQTSVAPPPVEVPSGVRIVVEYCEPCGFEATYLELASAVKEQVPGIEISRLG 60  
DB 1 MSGEPAFVSVVPPGVEAGSGVHIVVEYCKPGFEATYLELASAVKEQVPGIEISRLG 60  
QY 61 GTGAFEIENGQVFSKLENGGFPYVEKOLIEAIRRASNGETLEKITNSRPPCVIL 115  
DB 61 GTGAFEIENGQVFSKLENGGFPYVEKOLIEAIRRASNGEPVEKITNSRPPCVIL 115  
RESULT 3  
ID Q5ZIH7.CHICK PRELIMINARY; PRT; 126 AA.  
AC Q5ZIH7.  
DT 23-NOV-2004, integrated into UniProtKB/TrEMBL.  
DT 23-NOV-2004, sequence version 1.  
DE Hypothesized protein.  
DE Hypothetical protein.  
GN ORFNames=RCITB04.26b22;  
OS Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
OC Gallus.  
OX NCBI\_TaxID=9031;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=CB; TISSUE=Bursa;  
RA Caldwell R.B., Kierzek A.M., Arakawa H., Bezzubov Y., Zaim J.,  
RA Fiedler P., Kutter S., Blagodatski A., Kostovska D., Koter M.,  
RA Plachy J., Carninci P., Hayashizaki Y., Buerstedde J.M.,  
RT "Full-length cDNAs from chicken bursal lymphocytes to facilitate  
RT gene function analysis.";  
RL Genome Biol. 6:R6-R6(2005).  
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
CC Distributed under the Creative Commons Attribution-NoDerivs License  
CC EMBL; AJ270807; CAG32466.1; -; mRNA.  
DR GO; GO:0008430; F:selenium binding; IEA.  
DR GO; GO:0045454; P:cell redox homeostasis; IEA.  
DR InterPro; IPR011893; CXXU selWTH.  
DR TIGRFAMs; TIGR02174; CXXU\_selWTH; 1.  
KW Hypothetical protein.  
SQ SEQUENCE 126 AA; 13438 MW; 12B8498FD40DBB6D CRC64;  
Query Match 72.9%; Score 435.5; DB 2; Length 126;  
Best Local Similarity 72.2%; Pred. NO. 7.8e-35;  
Matches 91; Conservative 4; Mismatches 16; Indels 15; Gaps 2;  
QY 2 SGEPCQTSVAPPPVEVPSG-----VRIVVEYCEPCGFEATYLELASAVKEQ 49  
DB 4 SGGGAAAVG---TESEAGDGGFGSDSGSRRVHIVVEYCEPCGFEATYLELASAVRE 60  
QY 50 YFCIEIESRLGGTGAFEIENGQVFSKLENGGFPYVEKOLIEAIRRASNGETLEKITNSR 109  
DB 61 YPDIEIESRLGGTGAFEIENGQVFSKLENGGFPYVEKOLIEAIRRASNGEPVEKITNSR 120  
QY 110 PPCVIL 115  
||| |||











```

RL Submitted (JUL-2005) to the EMBL/GenBank/DDBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=TM1040;
RG US DOE Joint Genome Institute (JGI-PGF);
RA Copeland A., Lucas S., Lapidus A., Barry K., Detter C., Glavina T.,
RA Hammon N., Israni S., Pittluck S., Richardson P.;
RA Submitted (JUN-2005) to the EMBL/GenBank/DDBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AAFG02000004; EAN56920.1; -; Genomic_DNA.
DR GO; GO:0008430; P:selenium binding; IEA.
DR GO; GO:0045454; P:cell redox homeostasis; IEA.
DR InterPro; IPR011893; CXXU_selWTH.
DR InterPro; IPR007843; Seleno_W_rel.
DR Pfam; PF05169; Seleno_W_rel; 1.
DR TIGRFAMs; TIGR021174; CXXU_selWTH; 1.
SQ SEQUENCE 104 AA; 11690 MW; AB6A6793EBDB4AC1 CRC64;

Query Match 19.5%; Score 116.5; DB 2; Length 104;
Best Local Similarity 34.1%; Pred. No. 0.0019;
Matches 30; Conservative 16; Mismatches 35; Indels 7; Gaps 4;

QY 18 EPQSGVRIIVVECEPGCF--EATYL--ELASAVKEQYPCIE--IESRLGDTGAFETINGQ 72
DB 6 EPTTPHITITTCIGNWLLRAWSQELLSFTQEQLAGVTIVPGHIGT--FEISVEDQ 63

QY 73 LVFSKLENGFPYFKDLIEAIRASNGE 100
DB 64 LWERKDGDFDVKELKTRVDLINPE 91

RESULT 14
HSP6_HETGL
ID HSP6_HETGL STANDARD; PRT; 244 AA.
AC Q9BNT9;
DT 02-NOV-2001, integrated into UniProtKB/Swiss-Prot.
DT 01-JUN-2001, sequence version 1.
DT 07-FEB-2006, entry version 16.
DE Putative esophageal gland cell secretory protein 6 precursor.
GN Name=HSP6;
OS Heterodera glycines (Soybean cyst nematode worm).
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Tylenchida; Tylenchina;
OC Tylenchoidea; Heteroderidae; Heteroderinae; Heterodera.
OX NCBI_TaxID=51029;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA].
RX MEDLINE=21204659; PubMed=11310741;
RA Wang X., Allen R., Ding X., Goellner M., Maier T., de Boer J.M.,
RA Baum T.J., Husey R.S., Davis E.L.;
RT "Signal peptide-selection of cDNA cloned directly from the esophageal
RT gland cells of the soybean cyst nematode Heterodera glycines.";
RL Mol. Plant Microbe Interact. 14:536-544 (2001).
CC -!- SIMILARITY: Belongs to the SELT family.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AP273733; AAG21336.2; -; mRNA.
DR InterPro; IPR011893; CXXU_selWTH.
DR TIGRFAMs; TIGR021174; CXXU_selWTH; 1.
KW Hypothetical protein; Redox-active center; Signal.
FT SIGNAL 1 22 Potential.
FT CHAIN 23 244 Putative esophageal gland cell secretory
FT protein 6.
FT FTID=PRO_0000032296.
FT DISULFID 88 91 Redox-active (Potential).
FT SEQUENCE 244 AA; 27201 MW; DD94A7A590AA9143 CRC64;

Submitted (JUL-2005) to the EMBL/GenBank/DDBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=TM1040;
RG US DOE Joint Genome Institute (JGI-PGF);
RA Copeland A., Lucas S., Lapidus A., Barry K., Detter C., Glavina T.,
RA Hammon N., Israni S., Pittluck S., Richardson P.;
RA Submitted (JUN-2005) to the EMBL/GenBank/DDBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AAFG02000004; EAN56920.1; -; Genomic_DNA.
DR GO; GO:0008430; P:selenium binding; IEA.
DR GO; GO:0045454; P:cell redox homeostasis; IEA.
DR InterPro; IPR011893; CXXU_selWTH.
DR InterPro; IPR007843; Seleno_W_rel.
DR Pfam; PF05169; Seleno_W_rel; 1.
DR TIGRFAMs; TIGR021174; CXXU_selWTH; 1.
SQ SEQUENCE 104 AA; 11690 MW; AB6A6793EBDB4AC1 CRC64;

Query Match 19.2%; Score 114.5; DB 1; Length 244;
Best Local Similarity 22.5%; Pred. No. 0.008;
Matches 39; Conservative 15; Mismatches 30; Indels 89; Gaps 5;

QY 15 ERV---EPGSGVR-----IVVEYCEPCGFEATVLELASAVKEQYPCIE----- 54
DB 60 EVVTIREP-SGTKSEFKLPINPPVKFSCVCGRYAQVEQPAQLIREKYPQIDIHGENY 118
QY 55 ----- 54

DB 119 PPGILRTVGAQVIGVMKIALIVCVVSGRSPPTLGLTPTTFQWMLSNRLSAAMLFLFS 178

QY 55 --IESRLGDTGAFETINGQLVFSKLENGFPYFKDLIE-----AIRASNG 99
DB 179 NAIEGMLQSTGAFETIYESIRIWSKLESGRVSPPELFAQIDSHLAIRRGAG 231

RESULT 15
SEPW1_HUMAN
ID SEPW1_HUMAN STANDARD; PRT; 86 AA.
AC P63302; O15532; O19096; Q86TI9; Q96KMS;
DT 11-OCT-2004, integrated into UniProtKB/Swiss-Prot.
DT 11-OCT-2004, sequence version 1.
DT 07-FEB-2006, entry version 12.
DE Selenoprotein W.
GN Name=SEPW1; Synonyms=SELM;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Placenta;
RA Bellingham J., Gregory-Evans C.Y.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Skeletal muscle;
RX MEDLINE=97398151; PubMed=9256076; DOI=10.1016/S0378-1119(97)00113-3;
RA Gu Q.-P., Beilstein M.A., Vendeland S.C., Lugade A., Ream W.,
RA Whanger P.D.;
RT "Conserved features of selenocysteine insertion sequence (SECIS)
RT elements in selenoprotein W cDNAs from five species.";
RL Gene 193:187-196 (1997).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RA Bellingham J.;
RT "Genomic structure of human selenoprotein W (SEPW1).";
RL Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN [4]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC TISSUE=Brain, PNS, Testis, and Uterus;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.D., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Schetz T.E.,
RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullány S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
```

```
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: May be involved in a redox-related process. May play a
CC role in the myopathies of selenium deficiency (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasm.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL: AF015283; AAB69859.1; -; mRNA.
DR EMBL: U67171; AAC51665.1; -; mRNA.
DR EMBL: AF247455; AAL00898.1; -; Genomic DNA.
DR EMBL: AF247454; AAL00898.1; JOINED; Genomic DNA.
DR EMBL: BC000581; AAH00581.1; -; mRNA.
DR EMBL: BC032546; AAH32546.1; -; mRNA.
DR EMBL: BC039597; AAH39597.1; -; mRNA.
DR EMBL: BC047893; AAH47893.2; -; mRNA.
DR Ensembl: ENSG00000178980; Homo sapiens.
DR HGNC: HGNC:10752; SEPW1.
DR MIM: 603235; gene.
DR LinkHub: P63302; -.
DR GO: GO:0005737; Cytoplasm; NAS.
DR GO: GO:0005886; Cytoplasm membrane; ISS.
DR GO: GO:0016491; Peroxidoreductase activity; ISS.
DR InterPro: IPR011893; CXXU selWTH.
DR TIGRFAMs: TIGR02174; CXXU selWTH; 1.
KW Redox-active center; Selenium; Selenocysteine.
FT INIT_MET 0 0 By similarity.
FT CHAIN 1 86 Selenoprotein W.
FT SE CYS 12 12 Redox-active (By similarity).
FT DISULFID 9 12 Redox-active (By similarity).
FT SEQUENCE 86 AA; 9270 MW; 21297D8D1815F7E8 CRC64;
Query Match 18.8%; Score 112.5; DB 1; Length 86;
Best Local Similarity 33.3%; Pred. No. 0.0038;
Matches 27; Conservative 21; Mismatches 24; Indels 9; Gaps 4;
QY 23 VRIVVEYCPGFEATYLELASAVKQYPC-IEI--ESRLGGTGAFETIENGOLFVSKLE 79
DB 4 VRVW--YCGACGYKSKYLQKKLEDEFFGRLDICGEGTPQATGFFVWVAGKLIHSKKX 61
QY 80 NGGFPYVKD----LIEAIRRA 96
DB 62 GDGYVDTESKFLKLVAAIKAA 82
```

Search completed: December 6, 2006, 10:39:47  
Job time : 151 secs

**This Page Blank (uspio)**

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: December 6, 2006, 10:39:56 ; Search time 29 Seconds  
(without alignments)  
347.104 Million cell updates/sec

Title: US-09-824-787B-2

Perfect score: 597

Sequence: 1 MSCEPQTSVAPPPEVEPG.....ASNGETLEKITNSRPPCVIL 115

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 650591 seqs, 87530628 residues

Total number of hits satisfying chosen parameters: 650591

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:\*

- 1: /EMC\_Celerra\_SIDS3/ptodata/2/iaa/5 COMB.pdp.\*
- 2: /EMC\_Celerra\_SIDS3/ptodata/2/iaa/6 COMB.pdp.\*
- 3: /EMC\_Celerra\_SIDS3/ptodata/2/iaa/7 COMB.pdp.\*
- 4: /EMC\_Celerra\_SIDS3/ptodata/2/iaa/H COMB.pdp.\*
- 5: /EMC\_Celerra\_SIDS3/ptodata/2/iaa/PCTUS COMB.pdp.\*
- 6: /EMC\_Celerra\_SIDS3/ptodata/2/iaa/RE COMB.pdp.\*
- 7: /EMC\_Celerra\_SIDS3/ptodata/2/iaa/backfiles1.pdp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	403	67.5	90	2	US-09-513-999C-7234
2	96.5	16.2	88	1	US-08-807-043-3
3	96.5	16.2	88	1	US-09-127-289-3
4	96.5	16.2	88	2	US-09-393-448-3
5	96.5	16.2	93	1	US-08-807-043-1
6	96.5	16.2	93	1	US-09-127-289-1
7	96.5	16.2	93	2	US-09-393-448-1
8	87	14.6	216	2	US-09-270-767-46297
9	83	13.9	99	2	US-09-252-991A-17867
10	83	13.9	119	2	US-09-489-039A-10006
11	72	12.1	352	2	US-09-248-796A-17066
12	71.5	12.0	400	2	US-09-252-991A-22821
13	70	11.7	695	2	US-09-489-039A-14338
14	69	11.6	874	2	US-10-163-214-13
15	69	11.6	915	2	US-10-163-214-6
16	68.5	11.5	409	2	US-09-352-991A-29381
17	68	11.4	417	2	US-09-433-241A-10
18	68	11.4	735	2	US-09-115-704-2
19	68	11.4	735	2	US-09-780-115-2
20	68	11.4	909	2	US-10-163-214-2
21	67.5	11.3	1372	2	US-09-902-540-14099
22	67	11.2	239	2	US-09-720-318A-4
23	67	11.2	696	1	US-08-765-081-5
24	67	11.2	696	1	US-09-098-082-5
25	67	11.2	696	5	PCT-US95-06994-5
26	67	11.2	703	5	PCT-US95-06994-8

27 67 11.2 718 5 PCT-US95-06994-6 Sequence 6, Appli  
28 66.5 11.1 282 2 US-09-198-452A-739 Sequence 739, App  
29 66.5 11.1 287 2 US-09-438-185A-699 Sequence 699, App  
30 66.5 11.1 772 2 US-09-907-794A-339 Sequence 339, App  
31 66.5 11.1 772 2 US-09-905-125A-339 Sequence 339, App  
32 66.5 11.1 772 2 US-09-902-775A-339 Sequence 339, App  
33 66.5 11.1 772 2 US-09-906-700-339 Sequence 339, App  
34 66.5 11.1 772 2 US-09-903-603A-339 Sequence 339, App  
35 66.5 11.1 772 2 US-09-904-920A-339 Sequence 339, App  
36 66.5 11.1 772 2 US-09-909-064-339 Sequence 339, App  
37 66.5 11.1 772 2 US-09-905-381A-339 Sequence 339, App  
38 66.5 11.1 772 2 US-09-906-618-339 Sequence 339, App  
39 66.5 11.1 772 2 US-09-906-646-339 Sequence 339, App  
40 66.5 11.1 772 2 US-09-904-462-339 Sequence 339, App  
41 66.5 11.1 772 2 US-09-902-736A-339 Sequence 339, App  
42 66.5 11.1 772 2 US-09-906-722A-339 Sequence 339, App  
43 66.5 11.1 772 2 US-09-905-449-339 Sequence 339, App  
44 66.5 11.1 772 2 US-09-903-562B-339 Sequence 339, App  
45 66.5 11.1 772 2 US-09-906-679A-339 Sequence 339, App

## ALIGNMENTS

RESULT 1  
US-09-513-999C-7234  
; Sequence 7234, Application US/09513999C  
; Patent No. 6783961  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Duclert, A.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.  
; Patent No. 6783961  
; FILE REFERENCE: 59.US2.REG  
; CURRENT APPLICATION NUMBER: US/09/513,999C  
; CURRENT FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/122,487  
; PRIOR FILING DATE: 1999-02-26  
; NUMBER OF SEQ ID NOS: 36681  
; SOFTWARE: Patent.pm  
; SEQ ID NO 7234  
; LENGTH: 90  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: UNSURE  
; LOCATION: 18  
; OTHER INFORMATION: Xaa=Asp or Glu  
; FEATURE:  
; NAME/KEY: UNSURE  
; LOCATION: 46  
; OTHER INFORMATION: Xaa=Leu or Val  
; FEATURE:  
; NAME/KEY: UNSURE  
; LOCATION: 78  
; OTHER INFORMATION: Xaa=Ala or Gly  
US-09-513-999C-7234

Query Match 67.5%; Score 403; DB 2; Length 90;  
Best Local Similarity 87.6%; Pred. No. 4.5e-40;  
Matches 78; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

Qy 1 MSCEPQTSVAPPPEVEPGSGVRIYVEVCEPGFATYLELASAVKEQVPGIEISRLG 60  
Db 1 MSCEPQTSVAPPPEVEPGSGVRIYVEVCEPGFATYLELASAVKEQVPGIEISRLG 60  
Qy 61 GTGAFIEINGQLVFSKLENGGFPYEKDL 89  
Db 61 GTGAFIEINGQLVGLQXGEGFPYEKDV 89

RESULT 2





```

; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: BLADNOT03
; CLONE: 1599862
US-08-807-043-1

Query Match      16.2%; Score 96.5; DB 1; Length 93;
Best Local Similarity 32.8%; Pred. No. 0.001;
Matches 21; Conservative 18; Mismatches 20; Indels 5; Gaps 3;

QY 23 VRIVWCYPCGPEATYLELASAVKEQYPC-IBI--ESRLGGTGAFETIEINGOLVFVKLE 79
Db 11 VRVV--YCGAXGYKSYLQKKLEDEFFERLDICGEGTSQAXGFFEVWVAGKLIHKKX 68
QY 80 NGGF 83
Db 69 GDGY 72

RESULT 6
US-09-127-289-1
; Sequence 1, Application US/09127289
; Patent No. 5998371
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Goli, Surya K.
; TITLE OF INVENTION: NOVEL HUMAN SELENOPROTEIN
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/127,289
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA: 08/807,043
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0202 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 93 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: BLADNOT03
; CLONE: 1599862
US-09-127-289-1

Query Match      16.2%; Score 96.5; DB 1; Length 93;
Best Local Similarity 32.8%; Pred. No. 0.001;
Matches 21; Conservative 18; Mismatches 20; Indels 5; Gaps 3;

QY 23 VRIVWCYPCGPEATYLELASAVKEQYPC-IBI--ESRLGGTGAFETIEINGOLVFVKLE 79

```



; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 2709.2004001

; CURRENT APPLICATION NUMBER: US/09/489,039A

; CURRENT FILING DATE: 2000-01-27 1999-07-27

; PRIOR APPLICATION NUMBER: US 60/117,747

; PRIOR FILING DATE: 1999-01-29

; NUMBER OF SEQ ID NOS: 14342

; SEQ ID NO 10006

; LENGTH: 119

; TYPE: PRT

; ORGANISM: Klebsiella pneumoniae

US-09-489-039A-10006

Query Match

Best Local Similarity 13.9%; Score 83; DB 2; Length 119;

Matches 21; Conservative 17; Mismatches 39; Indels 6; Gaps 3;

QY 16 EYEGSGVRVVEYCPGCF--EATYL--ELASAVKEQYPGIEISRLGDTGAFIEING 71

DB 20 ETTMSKAAITTYCQCWMLRASWQAQELLHTFTSDIASVTLVPTGT--GIFTIDVG 77

QY 72 QLVFSKLENGPPYKDLIEAIR 94

DB 78 QQIWERKQDGGFPDAAELKRRVR 100

RESULT 11

US-09-248-796A-17066

; Sequence 17066, Application US/09248796A

; Patent No. 6747137

; GENERAL INFORMATION:

; APPLICANT: Keith Weinstein et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN

; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.132

; CURRENT APPLICATION NUMBER: US/09/248,796A

; CURRENT FILING DATE: 1999-02-12

; PRIOR APPLICATION NUMBER: US 60/074,725

; PRIOR FILING DATE: 1998-02-13

; PRIOR APPLICATION NUMBER: US 60/096,409

; PRIOR FILING DATE: 1998-08-13

; NUMBER OF SEQ ID NOS: 28208

; SEQ ID NO 17066

; LENGTH: 352

; TYPE: PRT

; ORGANISM: Candida albicans

US-09-248-796A-17066

Query Match

Best Local Similarity 12.1%; Score 72; DB 2; Length 352;

Matches 21; Conservative 14; Mismatches 32; Indels 18; Gaps 2;

QY 33 CGFEATYLELASAVKEQYPGIEISRLGDTGAFIEINGOLVFSKLE--NGG----- 82

DB 158 CAGVTYKALKTAELRAGQWVAISGAGGLGSLAVQAKAMGYRVLAIDGDKGEPVKS 217

QY 83 -----FPYKDLIEAIRASNG 99

DB 218 LGAETTFIDFTKEDVVEAVKATNG 242

RESULT 12

US-09-252-991A-22821

; Sequence 22821, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; CURRENT FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 2821

; LENGTH: 400

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-22821

Query Match

Best Local Similarity 12.0%; Score 71.5; DB 2; Length 400;

Matches 26; Conservative 10; Mismatches 33; Indels 19; Gaps 4;

QY 5 PGQTSVAPPEEVEP-----GSGVRVVEYCPGCFEATYLE--LASAVKEQYPG 52

DB 297 PATASPAAPPAPEAPAAAPVAGEGQGV-VKQFVADCTQVTDANGKVLVSALKRKGD 355

QY 53 IE-----IESRLGDTGAFIEINGQLV 74

DB 356 LELAGKAPLELRGLGFARGAQVSYNGQPV 383

RESULT 13

US-09-489-039A-14338

; Sequence 14338, Application US/09489039A

; Patent No. 6610836

; GENERAL INFORMATION:

; APPLICANT: Gary Breton et. al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLBBSIELLA

; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 2709.2004001

; CURRENT APPLICATION NUMBER: US/09/489,039A

; CURRENT FILING DATE: 2000-01-27 1999-07-27

; PRIOR APPLICATION NUMBER: US 60/117,747

; PRIOR FILING DATE: 1999-01-29

; NUMBER OF SEQ ID NOS: 14342

; SEQ ID NO 14338

; LENGTH: 695

; TYPE: PRT

; ORGANISM: Klebsiella pneumoniae

US-09-489-039A-14338

Query Match

Best Local Similarity 11.7%; Score 70; DB 2; Length 695;

Matches 18; Conservative 20; Mismatches 39; Indels 4; Gaps 1;

QY 29 YCEPCGFEATYLELASAVKEQYPGIEISRLGDTGAFIEINGOLVFSKLENGGFPYK 88

DB 531 YSGFVGMQT---LSNAGKADATGVELEAKWRFPAGMSWDINGVIRSEFTNDSELYHGN 586

QY 89 LIBAIRRASNGETLEKITNSR 109

DB 587 RVFPVPRYGAGSSVNGVIDTR 607

RESULT 14

US-10-163-214-13

; Sequence 13, Application US/10163214

; Patent No. 6849781

; GENERAL INFORMATION:

; APPLICANT: Allen, Stephen M.

; APPLICANT: Broglie, Karen E.

; APPLICANT: Butler, Karlene H.

; APPLICANT: Thorpe, Catherine J.

; TITLE OF INVENTION: Starch Synthase Isoform V

; FILE REFERENCE: BB1520 US NA

; CURRENT APPLICATION NUMBER: US/10/163,214

; CURRENT FILING DATE: 2002-06-05

; PRIOR APPLICATION NUMBER: 60/297,099

; PRIOR FILING DATE: 2001-06-08

; NUMBER OF SEQ ID NOS: 13

; SOFTWARE: Microsoft Office 97

```
/ SEQ ID NO 13
/ LENGTH: 874
/ TYPE: PRT
/ ORGANISM: Vigna unguiculata
US-10-163-214-13

Query Match      11.6%; Score 69; DB 2; Length 874;
Best Local Similarity 26.8%; Pred. No. 47;
Matches 22; Conservative 14; Mismatches 34; Indels 12; Gaps 3;

QY 21 SGVRIVVEYCEPCGPEATYLELASAVKEQYPGIEIESRLGGTGCAFEIEINGQLVFSKLEN 80
Db 754 SDMPFIIPSIFECG-----LTQMISMRYGAIPARTGGGLNDSVFDVDDDTIPSQFEN 806

QY 81 GGFY-----EKDLIEAIRASN 98
Db 807 -GFTFLNADKGINDAIVRAIN 827

RESULT 15
US-10-163-214-6
/ Sequence 6, Application US/10163214
/ Patent No. 6849781
/ GENERAL INFORMATION:
/ APPLICANT: Allen, Stephen M.
/ APPLICANT: Broglie, Karen E.
/ APPLICANT: Butler, Karlene H.
/ APPLICANT: Thorpe, Catherine J.
/ TITLE OF INVENTION: Starch Synthase Isoform V
/ FILE REFERENCE: BB1520 US NA
/ CURRENT APPLICATION NUMBER: US/10/163,214
/ CURRENT FILING DATE: 2002-06-05
/ PRIOR APPLICATION NUMBER: 60/297,099
/ PRIOR FILING DATE: 2001-06-08
/ NUMBER OF SEQ ID NOS: 13
/ SOFTWARE: Microsoft Office 97
/ SEQ ID NO 6
/ LENGTH: 915
/ TYPE: PRT
/ ORGANISM: Oryza sativa
US-10-163-214-6

Query Match      11.6%; Score 69; DB 2; Length 915;
Best Local Similarity 28.4%; Pred. No. 51;
Matches 23; Conservative 11; Mismatches 37; Indels 10; Gaps 2;

QY 21 SGVRIVVEYCEPCGPEATYLELASAVKEQYPGIEIESRLGGTGCAFEIEINGQLVFSKLEN 80
Db 797 SDMFIVPSMFECG-----LTQMTAMRYGSPVIVRTGGGLNDSVFDDEDTIPKELRN 849

QY 81 G---GPPYEKDLIEAIRASN 98
Db 850 GFTFVHPDEKALSGAMERAFN 870

Search completed: December 6, 2006, 10:40:50
Job time : 30 secs
```

Title: US-09-824-787B-2  
Perfect score: 597  
Sequence: 1 MSGEPQTSVAPPEVEPG.....ASNETLEKXTNSRPPCVIL 115  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5  
Searched: 2589679 seqs, 457216429 residues  
Total number of hits satisfying chosen parameters: 2589679  
Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq.8.\*  
1: Geneseq1980s.\*  
2: Geneseq1990s.\*  
3: Geneseq2000s.\*  
4: Geneseq2001s.\*  
5: Geneseq2002s.\*  
6: Geneseq2003as.\*  
7: Geneseq2003bs.\*  
8: Geneseq2004s.\*  
9: Geneseq2005s.\*  
10: Geneseq2006s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	597	100.0	115	4	AAG78997 Human C35
2	597	100.0	115	4	AAG78997 Human C35
3	597	100.0	115	5	ABP58560 Human sit
4	597	100.0	115	5	ABP43843 RIKEN 181
5	597	100.0	115	8	ADH13244 Human mal
6	597	100.0	115	8	ADK48916 Human bre
7	597	100.0	115	9	ADV60511 Human bre
8	597	100.0	115	9	ADK83703 Human C35
9	597	100.0	115	9	AEA15131 Human pol
10	597	100.0	115	9	AEA62447 Human C35
11	597	100.0	115	10	AEF13891 Human C35
12	597	100.0	124	5	ABP43055 Human ova
13	597	100.0	111	3	ABP43521 Human can
14	597	100.0	131	8	ADK49069 Human bre
15	597	100.0	149	9	AEA62539 Recombina
16	597	100.0	206	6	ABR47619 Breast ca
17	466.5	78.1	207	8	ADP84573 Human bre
18	466.5	78.1	208	8	ADP84574 Human bre
19	403	67.5	203	3	AG03153 Human sec
20	403	67.5	90	8	ADK49067 Human bre
21	292	48.9	75	8	ADK50220 Human car
22	291	48.7	65	8	ADK50219 Human car
23	290	48.6	65	8	ADK50217 Human car

24	290	48.6	75	8	ADK50218 Human car
25	289	48.4	64	8	ADK50201 Human car
26	289	48.4	73	8	ADK50214 Human car
27	289	48.4	73	8	ADK50202 Human car
28	288	48.2	64	8	ADK50187 Human car
29	288	48.2	64	8	ADK50213 Human car
30	288	48.2	73	8	ADK50188 Human car
31	288	48.2	75	8	ADK50196 Human car
32	287	48.1	64	8	ADK50209 Human car
33	287	48.1	64	8	ADK50215 Human car
34	287	48.1	73	8	ADK50210 Human car
35	287	48.1	73	8	ADK50216 Human car
36	286	47.9	64	8	ADK50227 Human car
37	286	47.9	64	8	ADK50221 Human car
38	286	47.9	66	8	ADK50225 Human car
39	286	47.9	73	8	ADK50222 Human car
40	286	47.9	73	8	ADK50182 Human car
41	286	47.9	73	8	ADK50228 Human car
42	285.5	47.8	77	8	ADK50226 Human car
43	285.5	47.8	85	8	ADK50206 Human car
44	285	47.7	64	8	ADK50229 Human car
45	285	47.7	65	8	ADK50189 Human car

ALIGNMENTS

RESULT 1	
AAG78997	
ID	AAG78997 standard; protein; 115 AA.
XX	
AC	AAG78997;
DT	
DT	22-JAN-2002 (first entry)
XX	
DE	Human C35, a tumour antigen.
XX	
KW	Human; C35; cytostatic; gene therapy; vaccine; tumour antigen;
KW	breast cancer; bladder cancer; tumour immunotherapy; chromosome 17q12.
XX	
OS	Homo sapiens.
XX	
PN	WO200174859-A2.
XX	
PD	11-OCT-2001.
XX	
PF	04-APR-2001; 2001WO-US010855.
XX	
PR	04-APR-2000; 2000US-0194463P.
XX	
PA	(UYRP ) UNIV ROCHESTER.
XX	
PI	Zauderer M, Evans EE, Borrello MA;
XX	
DR	WPI; 2001-626383/72.
DR	N-PSDB; AAI71785, AAI71793.
XX	
PT	Novel C35 polypeptides and C35 genes useful in immunogenic compositions
PT	and vaccines, for inducing antibody and cell-mediated immunity against
XX	target cells, such as tumor cells that express C35 gene.
XX	
PS	Claim 11; Fig 1; 331pp; English.
XX	
CC	The present sequence is human C35. C35 is a novel tumour antigen that is
CC	overexpressed in human breast and bladder carcinoma. C35 is thought to be
CC	a promising candidate for tumour immunotherapy, in immunogenic
CC	compositions and vaccines, to induce antibody and cell-mediated immunity
CC	against target cells such as tumour cells that express C35 genes. The C35
CC	gene aligns on human chromosome 17q12
XX	
SQ	Sequence 115 AA;
	Query Match 100.0%; Score 597; DB 4; Length 115;

Best Local Similarity 100.0%; Pred. No. 1.3e-60;  
Matches 115; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSEPGQTGVAPPPVEEVPSSGVRIVVEYCEPCGFATYLELASAVKEQYPGIEISRLG 60  
|  
Db 1 MSSEPGQTGVAPPPVEEVPSSGVRIVVEYCEPCGFATYLELASAVKEQYPGIEISRLG 60  
|

QY 61 GTCAFEIENGQLVFSKLENGGPFYEKDLEAIRRASNGETLEKITNSRPPCVIL 115  
|  
Db 61 GTCAFEIENGQLVFSKLENGGPFYEKDLEAIRRASNGETLEKITNSRPPCVIL 115  
|

RESULT 2  
AAG77870  
ID AAG77870 standard; protein; 115 AA.  
XX AAG77870;  
XX  
XX 08-MAY-2002 (first entry)  
XX Human C35 protein.  
XX

KW C35 protein; antigenic peptide; major histocompatibility complex;  
KW MHC-peptide complex; MHC; human; MHC class I alpha chain;  
KW beta-2 microglobulin; MHC class II alpha chain; MHC class II beta chain;  
KW vaccine; immune response modulation; hyperproliferative disorder;  
KW neoplasm; hypergammaglobulinaemia; viral infection; hepatitis;  
KW meningitis; bacterial infection; tuberculosis; gingivitis;  
KW parasitic infection; autoimmune disease; Hashimoto's disease;  
KW Graves' disease; rheumatoid arthritis; allergy; asthma; organ rejection;  
KW graft-versus-host disease; GVHD; breast cancer.  
XX  
XX Homo sapiens.  
OS  
XX WO200178768-A2.  
PN  
XX  
PD 25-OCT-2001.  
XX  
PF 12-APR-2001; 2001WO-US011912.  
XX  
XX 12-APR-2000; 2000US-0196472P.  
PR (UYRP ) UNIV ROCHESTER.  
XX  
PA Zauderer M, Smith ES;  
XX  
PI WFI; 2001-602927/68.  
DR N-PSDB; AAH77148.  
XX  
XX Novel compound comprising major histocompatibility complex-peptide complexes, used to modulate immune responses.

Example 16; Fig 7; 166pp; English.

The invention comprises a compound which contains one or more major histocompatibility complex (MHC)-peptide complexes, and an antibody specific for a cell surface marker. The complexes comprise an MHC class I alpha chain, a beta-2 microglobulin molecule and an antigenic peptide bound in the MHC groove. Alternatively, the complexes may comprise an MHC class II alpha chain, an MHC class II beta chain, and an antigenic peptide bound in the MHC groove. The complexes are linked to the carboxyl terminus of the antibody. The compounds of the invention can be used as a vaccine to modulate an immune response. The compounds of the invention are useful for treating: hyperproliferative disorders (e.g. neoplasms and hypergammaglobulinaemia); viral infections (e.g. hepatitis and meningitis); bacterial infections (e.g. tuberculosis and gingivitis); parasitic infections; autoimmune diseases (e.g. Hashimoto's disease, Graves' disease and rheumatoid arthritis); allergic reactions/conditions (e.g. asthma). The compounds of the invention may also be used in the treatment of organ rejection or graft-versus-host disease (GVHD). The present sequence represents C35 protein, this protein is differentially expressed in human breast cancer cells. In one embodiment of the invention, antigenic peptides derived from the C35 protein can be used to

QY 61 GTGAFIEINGQLVFSKLENGGPPYEKDLIEAIRRASNGETLEKITSRRPPCVIL 115  
 DB 61 GTGAFIEINGQLVFSKLENGGPPYEKDLIEAIRRASNGETLEKITSRRPPCVIL 115

RESULT 4  
 ABP43843  
 ID ABP43843 standard; protein; 115 AA.

AC ABP43843;

DT 26-FEB-2003 (first entry)

DE RIKEN 1810046J19 protein.

XX Neuroprotective; immunomodulator; cancer; chromosome 4q13-q21;  
 KW cytostatic; anti-inflammatory; gene therapy; nutritional supplement;  
 KW wound; burn; ulcer; Alzheimer's disease; Huntington's disease;  
 KW amyotrophic lateral sclerosis; autoimmune disorder; inflammation;  
 KW vulnery.

OS Homo sapiens.

PN WO200231111-A2.

XX 18-APR-2002.

PF 11-OCT-2001; 2001WO-US027760.

XX 12-OCT-2000; 2000US-00687527.

XX (HYSE-) HYSEQ INC.

PI Tang YT, Liu C, Zhou P, Aundi V, Zhang J, Zhao QA, Ren F;

PI Xue AJ, Yang Y, Wehrman T, Drmanac RT;

DR WPI: 2002-426278/45.

DR N-PSDB; ABQ61087.

XX New polypeptides and their encoded proteins, useful as nutritional  
 PT sources or supplements, or in gene therapy, particularly for treating  
 PT wounds, Alzheimer's disease, amyotrophic lateral sclerosis, cancer or  
 PT inflammation.

XX Claim 20; SEQ ID # 746; 357pp + Sequence Listing; English.

XX The invention relates to 446 newly isolated polynucleotide sequences. The  
 CC activity of polynucleotides of the invention may be described as,  
 CC vulnery, neuroprotective, immunomodulator, cytostatic and anti-  
 CC inflammatory. Compositions comprising nucleic acids of the invention are  
 CC useful for treating a mammalian subject, or as nutritional sources or  
 CC supplements. These are useful in gene therapy, particularly for treating  
 CC wounds, burns or ulcers, Alzheimer's disease, Huntington's disease,  
 CC amyotrophic lateral sclerosis, autoimmune disorders, cancer or  
 CC inflammation. The nucleic acids and polypeptides are also useful in  
 CC diagnostic and research methods. The sequences given in records ABP43544-  
 CC ABP43989 represent polypeptides encoded by polynucleotides of the  
 CC invention. NOTE: The sequence data for this patent did not form part of  
 CC the printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 115 AA;

Query Match 100.0%; Score 597; DB 5; Length 115;

Best Local Similarity 100.0%; Pred. No. 1.3e-60;

Matches 115; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSCEPGQTSVAPPEEVEPGSGVRIIVVEYCEPCGFEATYLELASAVKEQVPGIEISRLG 60

DB 1 MSCEPGQTSVAPPEEVEPGSGVRIIVVEYCEPCGFEATYLELASAVKEQVPGIEISRLG 60

QY 61 GTGAFIEINGQLVFSKLENGGPPYEKDLIEAIRRASNGETLEKITSRRPPCVIL 115

DB 61 GTGAFIEINGQLVFSKLENGGPPYEKDLIEAIRRASNGETLEKITSRRPPCVIL 115

RESULT 5

ADH13244

ID ADH13244 standard; protein; 115 AA.

XX ADH13244;

DT 11-MAR-2004 (first entry)

XX Human malignant neoplasia-related protein SeqID93.

XX malignant neoplasia; cytostatic; breast cancer; ovarian cancer;  
 KW gastric cancer; colon cancer; oesophageal cancer; mesenchymal cancer;  
 KW bladder cancer; non-small cell lung cancer; human.

OS Homo sapiens.

PN EP1365034-A2.

XX 26-NOV-2003.

XX 09-MAY-2003; 2003EP-00010447.

XX 21-MAY-2002; 2002EP-00010291.

XX 13-FEB-2003; 2003EP-00003112.

XX (FARB ) BAYER AG.

XX Wirtz R, Munnes M, Kallabis H;

XX WPI; 2004-073279/08.

XX N-PSDB; ADH13221.

XX Predicting, diagnosing or prognosing malignant neoplasia by detecting at  
 PT least two markers, where the markers are genes from one or more  
 PT chromosomal regions altered in malignant neoplasia.

XX Claim 11; SEQ ID NO 93; 267pp; English.

XX This invention relates to a novel method for the prediction, diagnosis,  
 CC or prognosis of malignant neoplasia by the detection of at least two  
 CC markers. The invention may also be useful for the development of  
 CC cytostatic compounds through the regulation of the expression of a gene  
 CC or activity of a protein associated with malignant neoplasia. The method  
 CC is useful for prediction, diagnosis or prognosis of malignant neoplasia  
 CC such as breast cancer, ovarian cancer, gastric cancer, colon cancer,  
 CC oesophageal cancer, mesenchymal cancer, bladder cancer or non-small cell  
 CC lung cancer. The polynucleotides and polypeptides defined in the  
 CC specification, antisense polynucleotides targeting the polynucleotides,  
 CC antibodies targeting either one of the polynucleotides or polypeptides,  
 CC and compounds identified by the screening methods are useful for  
 CC preventing or treating malignant neoplasia. The disease treated is  
 CC preferably breast cancer. The present sequence is that of a human  
 CC malignant neoplasia-related protein which may be used in the method of  
 CC the invention.

XX Sequence 115 AA;

Query Match 100.0%; Score 597; DB 8; Length 115;

Best Local Similarity 100.0%; Pred. No. 1.3e-60;

Matches 115; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSCEPGQTSVAPPEEVEPGSGVRIIVVEYCEPCGFEATYLELASAVKEQVPGIEISRLG 60

DB 1 MSCEPGQTSVAPPEEVEPGSGVRIIVVEYCEPCGFEATYLELASAVKEQVPGIEISRLG 60

QY 61 GTGAFIEINGQLVFSKLENGGPPYEKDLIEAIRRASNGETLEKITSRRPPCVIL 115

DB 61 GTGAFIEINGQLVFSKLENGGPPYEKDLIEAIRRASNGETLEKITSRRPPCVIL 115

RESULT 6  
 ADK48916  
 ID ADK48916 standard; protein; 115 AA.  
 XX  
 AC ADK48916;  
 XX  
 DT 04-NOV-2004 (first entry)  
 XX  
 DE Human breast/bladder carcinoma-related C35 protein.  
 XX  
 KW C35 epitope; cytostatic; vaccine; tumour; breast; bladder carcinoma;  
 KW human; C35.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO2003104428-A2.  
 XX  
 PD 18-DEC-2003.  
 XX  
 PF 10-JUN-2003; 2003WO-US018252.  
 XX  
 PR 10-JUN-2002; 2002US-0386738P.  
 PR 11-DEC-2002; 2002US-0432241P.  
 PR 23-APR-2003; 2003US-0464550P.  
 XX  
 PA (VACC-) VACCINEX INC.  
 PA (UYRP-) UNIV ROCHESTER.  
 XX  
 PI Zauderer M, Evans EE, Borrello MA;  
 XX  
 DR WPI; 2004-062349/06.  
 DR N-PSDB; ADK48915.  
 XX  
 PT Novel C35 polypeptide useful for formulation of immunogenic composition  
 PT to induce antibodies and cell-mediated immunity against tumor cells.  
 XX  
 PS Claim 1; SEQ ID NO 2; 626pp; English.  
 XX  
 CC The invention relates to a novel isolated polypeptide comprising or  
 CC consisting of two or more C35 peptide epitopes. The polypeptide of the  
 CC invention demonstrates cytostatic activity and may be useful for the  
 CC formulation of an immunogenic composition, such as a vaccine, to induce  
 CC antibodies and cell-mediated immunity against target cells such as tumour  
 CC cells. Furthermore, the polypeptide and its analogues may be useful as  
 CC prognostic markers for carcinoma, such as human breast or bladder  
 CC carcinoma. The current sequence is that of human breast/bladder carcinoma  
 CC -related C35 protein of the invention.  
 XX  
 SQ Sequence 115 AA;  
 Query Match 100.0%; Score 597; DB 8; Length 115;  
 Best Local Similarity 100.0%; Pred. No. 1.3e-60;  
 Matches 115; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MSGEGQTSVAPPPPEVEPGSGVRIIVVEYCEPCGFEATYLELASAVKEQYPGIEISRLG 60  
 DB 1 MSGEGQTSVAPPPPEVEPGSGVRIIVVEYCEPCGFEATYLELASAVKEQYPGIEISRLG 60  
 QY 61 GTGAPEIENGQLVFSKLENGGFPPEKDLIEAIRRASNGETLEKITSRPPCVIL 115  
 DB 61 GTGAPEIENGQLVFSKLENGGFPPEKDLIEAIRRASNGETLEKITSRPPCVIL 115  
 RESULT 7  
 ADV60511  
 ID ADV60511 standard; protein; 115 AA.  
 XX  
 AC ADV60511;  
 XX  
 DT 24-FEB-2005 (first entry)  
 XX  
 DE Human breast cancer staging marker protein Seq 86.

XX breast tumor; cytostatic; diagnosis; prognosis; tumour marker; cancer.  
 KW Homo sapiens.  
 OS WO2004106495-A2.  
 XX  
 PN 09-DEC-2004.  
 XX  
 PD 26-MAY-2004; 2004WO-US016793.  
 XX  
 PF 29-MAY-2003; 2003US-0474281P.  
 PR 23-MAR-2004; 2004US-0555557P.  
 XX  
 PA (MILL-) MILLENNIUM PHARM INC.  
 XX  
 PI Mohahan JE, Hoersch S, Anderson DL, Endege WO, Ford D, Glatt K;  
 PI Gorbacheva BO, Kamatkar S, Xu Y, Gannavarapu M, Zhao X, Schlegel R;  
 PI Mertens M;  
 XX  
 DR WPI; 2005-039718/04.  
 DR N-PSDB; ADV60510.  
 XX  
 PT Novel marker protein, useful for assessing whether patient is afflicted  
 PT with breast cancer, for assessing efficacy of therapy for inhibiting  
 PT breast cancer, for assessing breast cell carcinogenic potential of test  
 PT composition.  
 XX  
 PS Disclosure; SEQ ID NO 86; 279pp; English.  
 XX  
 CC This invention relates to a novel secreted human marker proteins (Mks)  
 CC and the encoding nucleic acid molecules thereof. Specifically, it refers  
 CC to a method for assessing whether a patient has breast cancer that has  
 CC metastasized or is likely to metastasize. The present invention describes  
 CC determining the level of expression of a marker protein in the patient  
 CC sample and comparing this to the level from a control subject having a  
 CC non-metastasized breast tumor or no breast tumor. Furthermore it provides  
 CC a screening method for assessing the suitability of one or more test  
 CC compounds at inhibiting breast cancer in a patient by inhibiting the  
 CC expression of the marker proteins as given in the specification. As such,  
 CC it is useful for predicting the clinical outcome of a breast cancer  
 CC patient, for monitoring progression of the disease and for assessing the  
 CC breast cell carcinogenic potential of a test composition. Accordingly,  
 CC pharmaceutical compositions derived thereof exhibit cytostatic  
 CC activities. This polypeptide is a human breast cancer staging marker  
 CC protein of the invention.  
 XX  
 SQ Sequence 115 AA;  
 Query Match 100.0%; Score 597; DB 9; Length 115;  
 Best Local Similarity 100.0%; Pred. No. 1.3e-60;  
 Matches 115; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MSGEGQTSVAPPPPEVEPGSGVRIIVVEYCEPCGFEATYLELASAVKEQYPGIEISRLG 60  
 DB 1 MSGEGQTSVAPPPPEVEPGSGVRIIVVEYCEPCGFEATYLELASAVKEQYPGIEISRLG 60  
 QY 61 GTGAPEIENGQLVFSKLENGGFPPEKDLIEAIRRASNGETLEKITSRPPCVIL 115  
 DB 61 GTGAPEIENGQLVFSKLENGGFPPEKDLIEAIRRASNGETLEKITSRPPCVIL 115  
 RESULT 8  
 ADX83703  
 ID ADX83703 standard; protein; 115 AA.  
 XX  
 AC ADX83703;  
 XX  
 DT 05-MAY-2005 (first entry)  
 XX  
 DE Human C35 antigenic protein, seqid:2.  
 XX  
 KW Vaccine; immune modulation; delivery mechanism; microglobulin;



KW chemotherapy; gene therapy; adoptive immunotherapy;  
KW major histocompatibility complex; cancer; cytostatic; neoplasm;  
KW infectious disease; antimicrobial; infection; autoimmune disease;  
KW immunosuppressive; immune disorder; allergy; antiallergic; C35 DNA.  
XX  
OS Homo sapiens.  
XX  
PN US2005042218-A1.  
XX  
PD 24-FEB-2005.  
XX  
XX 09-JUL-2004; 2004US-00887230.  
XX  
XX 10-JUL-2003; 2003US-0485716P.  
PR 22-OCT-2003; 2003US-0513043P.  
XX  
XX (VACC-) VACCINEX INC.  
PA  
XX  
XX Zauderer M;  
PI  
XX WPI; 2005-180769/19.  
DR N-PSDB; ADX83702.  
XX  
XX New compounds comprising major histocompatibility complex I-peptide  
PT -antibody conjugates with modified beta2-microglobulin, useful for  
PT modulating immune responses or for treating or preventing e.g. cancer or  
PT infections.  
XX  
XX Disclosure; SEQ ID NO 2; 65pp; English.  
XX  
XX The present invention provides a novel targeted vaccine delivery system  
CC comprising one or more peptide-major histocompatibility complex (MHC)  
CC Class I complexes linked through the beta2-microglobulin molecule to an  
CC antibody which is specific for a cell surface marker. The invention is  
CC useful for modulating an immune response and for preventing and treating  
CC cancer, infectious diseases, autoimmune diseases and allergies. The  
CC present sequence is human C35 antigenic protein.  
XX  
XX Sequence 115 AA;  
SQ  
Query Match 100.0%; Score 597; DB 9; Length 115;  
Best Local Similarity 100.0%; Pred. No. 1.3e-60;  
Matches 115; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MSCEPGQTSVAPPEEVEPGSGVRIVVEYCEPCGFEATYLELASAVKEQYPGIEISRLG 60  
Db 1 MSCEPGQTSVAPPEEVEPGSGVRIVVEYCEPCGFEATYLELASAVKEQYPGIEISRLG 60  
QY 61 GTGAFIEINGQLVFSKLENGGFPYKDLIEAIRRASNGETLEKITSRPPCVIL 115  
Db 61 GTGAFIEINGQLVFSKLENGGFPYKDLIEAIRRASNGETLEKITSRPPCVIL 115  
RESULT 9  
AEA15131  
ID AEA15131 standard; protein; 115 AA.  
XX  
XX AEA15131;  
AC  
XX  
XX 28-JUL-2005 (first entry)  
DT  
XX  
DE Human polypeptide #44.  
XX  
KW Diagnosis; prognosis; cancer; breast tumor; ovary tumor; stomach tumor;  
KW colon tumor; esophagus tumor; bladder tumor; non-small-cell lung cancer;  
KW cytostatic; neoplasm.  
XX  
XX Homo sapiens.  
OS  
XX  
XX WO2005047534-A2.  
PN  
XX  
XX 26-MAY-2005.  
PD  
XX

PF 15-OCT-2004; 2004WO-BF011599.  
XX  
XX 28-OCT-2003; 2003EP-00024565.  
XX  
XX (FARB ) BAYER HEALTHCARE AG.  
XX  
XX Wirtz R, Munnes M;  
PI  
XX WPI; 2005-372393/38.  
DR N-PSDB; AEA15108.  
XX  
XX Predicting a response to cancer treatment by detecting at least 2  
PT markers, which are genes or genomic nucleic acid sequences that are  
PT located on one chromosomal region, which is altered in malignant  
PT neoplasia.  
XX  
XX Claim 8; SEQ ID NO 93; 464pp; English.  
XX  
XX The invention relates to a method of predicting response to cancer  
CC treatment comprising detection of at least 2 markers, where the markers  
CC are genes and fragments or genomic nucleic acid sequences that are  
CC located on one chromosomal region, which is altered in malignant  
CC neoplasia. The invention also relates to a method for the prediction,  
CC diagnosis or prognosis of malignant neoplasia, methods for detecting  
CC deregulations in malignant neoplasia and breast cancer, a method of  
CC determining the phenotype of a cell or tissue, a method for identifying  
CC genomic regions which are altered on the chromosomal level and encode  
CC genes that are linked by function and are differentially expressed in  
CC malignant neoplasia and breast cancer, methods of screening for agents  
CC which regulate the activity of a polypeptide or a polynucleotide and  
CC antibodies that specifically bind to a full length or partial  
CC polypeptide. The method is useful for predicting response to cancer  
CC treatment. The methods and compositions are useful for predicting,  
CC diagnosing, prognosing, preventing or treating malignant neoplasia  
CC including breast cancer, ovarian cancer, gastric cancer, colon cancer,  
CC esophageal cancer, mesenchymal cancer, bladder cancer or non-small-cell  
CC lung cancer. This sequence represents a human polypeptide used in the  
CC scope of the invention.  
XX  
XX Sequence 115 AA;  
SQ  
Query Match 100.0%; Score 597; DB 9; Length 115;  
Best Local Similarity 100.0%; Pred. No. 1.3e-60;  
Matches 115; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MSCEPGQTSVAPPEEVEPGSGVRIVVEYCEPCGFEATYLELASAVKEQYPGIEISRLG 60  
Db 1 MSCEPGQTSVAPPEEVEPGSGVRIVVEYCEPCGFEATYLELASAVKEQYPGIEISRLG 60  
QY 61 GTGAFIEINGQLVFSKLENGGFPYKDLIEAIRRASNGETLEKITSRPPCVIL 115  
Db 61 GTGAFIEINGQLVFSKLENGGFPYKDLIEAIRRASNGETLEKITSRPPCVIL 115  
RESULT 10  
AEA62447  
ID AEA62447 standard; protein; 115 AA.  
XX  
XX AEA62447;  
AC  
XX  
XX 25-AUG-2005 (first entry)  
DT  
XX  
XX Human C35 antigen.  
DE  
XX  
XX C35 antigen; antigen; cell growth; cancer; cytostatic; apoptosis;  
KW immunotherapy; hyperproliferation.  
XX  
XX Homo sapiens.  
OS  
XX  
XX WO200505936-A2.  
PN  
XX  
XX 23-JUN-2005.  
PD  
XX

PF 06-DEC-2004; 2004WO-US040573.  
 XX  
 PR 04-DEC-2003; 2003US-0526572P.  
 PR 23-DEC-2003; 2003US-0531688P.  
 XX  
 PA (VACC-) VACCINEX INC.  
 XX  
 PI Evans EE, Paris MJ, Sahasrabudhe DM, Smith ES, Zauderer M;  
 XX WPI; 2005-458501/46.  
 DR N-PSDB; AEA62446.  
 DR  
 XX  
 PT Killing cancer cells, by administering apoptosis-inducing therapy and  
 PT administering antibody specific for intracellular, cancer-associated  
 PT protein other than C35, or antibody specific for C35.  
 XX  
 PS Disclosure; SEQ ID NO 2; 255pp; English.  
 XX  
 CC The invention relates to killing (M1) cancer cells, comprising  
 CC administering an apoptosis-inducing therapy to cancer cells, and  
 CC administering to the cells an antibody specific for an intracellular,  
 CC cancer-associated protein, provided that the protein is not C35 antigen,  
 CC where protein becomes exposed on the cell surface in cells undergoing  
 CC apoptosis, where the antibody is conjugated to or complexed with a toxin.  
 CC The non-C35 antigen protein is a prenylated protein. Also included are an  
 CC isolated antibody (I) specific for C35 (chosen from an antibody  
 CC comprising the VH region encoded by clone 1B3G, the VL region encoded by  
 CC clone 1B3K, the VH region encoded by clone 1P2G, the VL region encoded by  
 CC clone 1P2K, the VH region encoded by clone H0009, the VL region encoded  
 CC by clone L0010, an antibody comprising at least one of CDR1 or CDR2 of  
 CC the VH region encoded by AEA62495, an antibody comprising at least one of  
 CC CDR1 or CDR2 of the VH region encoded by AEA62499, an antibody comprising  
 CC at least one of CDR1, CDR2, or CDR3 of the VL region encoded by  
 CC AEA62497, a chimeric antibody, or a humanized antibody), a polynucleotide  
 CC encoding the antibody, a vector comprising the polynucleotide, a host  
 CC cell comprising the vector and a composition comprising the antibody and  
 CC a carrier. The method is useful for killing cancer cells in a mammal  
 CC preferably human in need of eradication of smaller tumors and/or  
 CC micrometastases, or in need of cancer treatment for C35-associated cancer  
 CC chosen from breast cancer, ovarian cancer, bladder cancer, lung cancer,  
 CC prostate cancer, pancreatic cancer, colon cancer, melanoma and other  
 CC hyperproliferative disorders. The antibody is useful for detecting,  
 CC diagnosing or monitoring C35-associated cancers. The antibody comprises a  
 CC chimeric antibody comprising human immunoglobulin constant regions fused  
 CC to the variable regions of mouse anti-35 antibodies (named 1F2, 1B3, MAb  
 CC 165 and MAb 171). The present sequence represents the human C35 antigen.  
 XX  
 SQ Sequence 115 AA;  
 Query Match 100.0%; Score 597; DB 9; Length 115;  
 Best Local Similarity 100.0%; Pred. No. 1.3e-60;  
 Matches 115; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MSGEPQTSVAPPPPEVPGSGVRIIVVEYCEPCGFEATYLELASAVKEQYPGIEISRLG 60  
 DB 1 MSGEPQTSVAPPPPEVPGSGVRIIVVEYCEPCGFEATYLELASAVKEQYPGIEISRLG 60  
 QY 61 GTGAPEIENGQLVFSKLENGGPPYEKDLIEAIRRASNGETLEKITNSRPPCVIL 115  
 DB 61 GTGAPEIENGQLVFSKLENGGPPYEKDLIEAIRRASNGETLEKITNSRPPCVIL 115  
 RESULT 11  
 AEF13891  
 ID AEF13891 standard; protein; 115 AA.  
 XX  
 AC AEF13891;  
 XX  
 DT 09-MAR-2006 (first entry)  
 XX  
 DE Human C35 protein.  
 XX  
 XX Tumor marker; diagnosis; antigen; cancer; cytostatic; breast tumor;

KW colorectal tumor; prostate tumor; pancreas tumor; lung tumor;  
 KW ovary tumor; renal tumor; melanoma; immunotherapy.  
 OS Homo sapiens.  
 XX  
 PN WO2006002114-A2.  
 XX  
 PD 05-JAN-2006.  
 XX  
 PF 17-JUN-2005; 2005WO-US021836.  
 XX  
 PR 17-JUN-2004; 2004US-0580969P.  
 XX  
 PA (MANN-) MANNKIND CORP.  
 XX  
 PI Chiang C, Simard JUL;  
 XX  
 DR WPI; 2006-090202/09.  
 DR N-PSDB; AEF13911.  
 DR REFSEQ; NP\_115715.  
 XX  
 PT Matching a cancer condition in a patient with an immunotherapeutic agent  
 PT comprises assaying tumor tissue for at least two expressed tumor-  
 PT associated antigens.  
 XX  
 PS Disclosure; SEQ ID NO 11; 104pp; English.  
 XX  
 CC The invention relates to matching a cancer condition in a patient with an  
 CC immunotherapeutic agent or immunotherapeutic regimen comprising assaying  
 CC tumor tissue of the patient for two or more expressed tumor-associated  
 CC antigens (TuAAs) in a preselected panel, to develop an antigen profile  
 CC for the tumor. Also included is a method of confirming a cancer  
 CC diagnosis. At least one of the TuAAs is a cancer testis antigen, tissue-  
 CC specific antigen, oncofetal antigen, differentiation antigen, growth  
 CC factor, growth factor receptor, adhesion factor, signal transduction  
 CC product, transcription factor, oncogene product, tumor suppressor gene  
 CC product, or a microbial antigen. The preselected panel comprises two or  
 CC more antigens selected from an SSX protein, SSX-2, SSX-4, a MAGB protein,  
 CC MAGB-1, MAGB-3, PRAME, NY-ESO-1, LAGE, PSMA, SCP-1, melan-A/MART-1,  
 CC or tyrosinase, and where the cancer condition is carcinoma selected from  
 CC breast, colorectal, prostate, pancreatic, lung, ovarian, renal cell, or  
 CC melanocyte. The methods are useful for matching a cancer condition in a  
 CC patient with an immunotherapeutic agent or immunotherapeutic regimen and  
 CC for confirming a cancer diagnosis, where the cancer condition is  
 CC carcinoma selected from breast, colorectal, prostate, pancreatic, lung,  
 CC ovarian, renal cell, or melanocyte. The present sequence is a human tumor  
 CC -associated antigen used in the method of the invention, chosen from the  
 CC list above.  
 XX  
 SQ Sequence 115 AA;  
 Query Match 100.0%; Score 597; DB 10; Length 115;  
 Best Local Similarity 100.0%; Pred. No. 1.3e-60;  
 Matches 115; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MSGEPQTSVAPPPPEVPGSGVRIIVVEYCEPCGFEATYLELASAVKEQYPGIEISRLG 60  
 DB 1 MSGEPQTSVAPPPPEVPGSGVRIIVVEYCEPCGFEATYLELASAVKEQYPGIEISRLG 60  
 QY 61 GTGAPEIENGQLVFSKLENGGPPYEKDLIEAIRRASNGETLEKITNSRPPCVIL 115  
 DB 61 GTGAPEIENGQLVFSKLENGGPPYEKDLIEAIRRASNGETLEKITNSRPPCVIL 115  
 RESULT 12  
 ABP43055  
 ID ABP43055 standard; protein; 124 AA.  
 XX  
 AC ABP43055;  
 XX  
 DT 22-AUG-2002 (first entry)  
 XX  
 DE Human ovarian antigen HVCAA68, SEQ ID NO:4187.

XX KW Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;  
 KW ovarian cancer; breast; cancer; tumour; reproductive system disorder;  
 KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;  
 KW PCOS; ovarian cyst; dysmenorrhea; endocrine disorder; infection;  
 KW inflammatory condition; immune disorder; blood disorder;  
 KW cardiovascular disorder; respiratory system disorder; neurological disorder;  
 KW gastrointestinal disorder; urinary system disorder; drug screening;  
 KW gene therapy; chromosome mapping; forensic analysis;  
 KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;  
 KW antiinflammatory; gynaecological; reproductive.  
 XX OS Homo sapiens.  
 XX WO200200677-A1.  
 XX 03-JAN-2002.  
 XX 07-JUN-2001; 2001WO-US018569.  
 XX 07-JUN-2000; 2000US-0209467P.  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 XX Birse CE, Rosen CA;  
 PI WPI: 2002-147878/19.  
 DR N-PSDB; ABQ56132.  
 XX Isolated nucleic acid molecules encoding novel ovarian polypeptides,  
 PT useful in the prevention, treatment and diagnosis of cancer (e.g. ovarian  
 PT cancer), immune disorders, cardiovascular disorders and neurological  
 PT diseases.  
 XX Claim 11; SEQ ID NO 4187; 2922pp; English.  
 XX The invention relates to 2175 novel human ovarian antigens (ABP41054-  
 CC ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also  
 CC encompasses polypeptides 90% identical and polynucleotides 95% identical  
 CC to the sequences of the invention. The invention additionally relates to  
 CC recombinant vectors and host cells comprising human ovarian antigen  
 CC polynucleotides, antibodies against human ovarian antigens, and the use  
 CC of ovarian antigen polynucleotides and polypeptides in diagnosing,  
 CC treating, prognosing or preventing various ovary and/or breast-related  
 CC disorders. Such conditions include ovarian cancer and breast cancer, and  
 CC metastatic tumours of ovarian or breast origin, reproductive system  
 CC disorders (e.g., infertility, disorders of pregnancy, anovulation,  
 CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine  
 CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic  
 CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and  
 CC vaginitis), immune disorders (e.g., congenital and acquired  
 CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),  
 CC blood-related disorders (e.g., anaemia), cardiovascular disorders,  
 CC respiratory disorders, neurological disorders, gastrointestinal disorders  
 CC and urinary system disorders. Ovarian antigen polypeptides and  
 CC polynucleotides may also be used in screening for compounds which  
 CC modulate ovarian antigen expression or activity. The polynucleotides may  
 CC further be used for gene therapy, chromosome mapping, in the  
 CC identification of individuals and in forensic analysis, and the  
 CC polypeptides may be used as food additives or to prepare antibodies  
 CC useful in disease diagnosis, drug targeting and phenotyping. The present  
 CC sequence represents a human ovarian antigen of the invention. Note: The  
 CC sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences  
 XX Sequence 124 AA;  
 SQ Query Match 100.0%; Score 597; DB 5; Length 124;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-60;  
 Matches 115; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 1 MSGEPCQTSVAPPPEVPGSGVRIIVVEYCEPCGFEATYLELASAVKEQVPGIEISRLG 60

Db 10 MSGEPCQTSVAPPPEVPGSGVRIIVVEYCEPCGFEATYLELASAVKEQVPGIEISRLG 69  
 QY 61 GTGAFETIEINGOLVSKLENGGPPYEKDLIEAIRRASNGETLEKITNSRPPCVIL 115  
 Db 70 GTGAFETIEINGOLVSKLENGGPPYEKDLIEAIRRASNGETLEKITNSRPPCVIL 124  
 RESULT 13  
 AAB43521  
 ID AAB43521 standard; protein; 131 AA.  
 XX AAB43521;  
 AC  
 XX 08-FEB-2001 (first entry)  
 DT  
 XX Human cancer associated protein sequence SEQ ID NO:966.  
 DE  
 XX Human; cancer associated gene; cancer antigen; detection; cancer;  
 KW diagnosis; cytostatic; proliferative; vulnery; immunomodulatory;  
 KW antidiabetic; antiasthmatic; antirheumatic; antiarthritic; antiviral;  
 KW antiinflammatory; antithyroid; antiallergic; antibacterial; cardiant;  
 KW dermatological; neuroprotective; thrombolytic; coagulant; nootropic;  
 KW vasotropic; antipsoriatic; antiangiogenic; gene therapy; inflammation;  
 KW immune disorder; haematopoietic cell disorder; autoimmune disorder;  
 KW allergic reaction; graft versus host disease; organ rejection;  
 KW haemostatic; thrombolytic; cardiovascular disorder; infection;  
 KW neurological disease; drug screening.  
 XX Homo sapiens.  
 OS  
 XX WO200055350-A1.  
 PN  
 XX 21-SEP-2000.  
 PD  
 XX 08-MAR-2000; 2000WO-US005882.  
 PF  
 XX 12-MAR-1999; 99US-0124270P.  
 PR  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 PA  
 XX Rosen CA, Ruben SM;  
 PI WPI: 2000-587533/55.  
 DR N-PSDB; AAC77730.  
 DR  
 XX Novel isolated nucleic acids comprising sequences encoding peptides  
 PT useful for treating or diagnosing e.g. cancer.  
 PT  
 XX Claim 11; Page 1534-1535; 2352pp; English.  
 PS  
 XX AAC77607 to AAC78448 encode the human cancer associated proteins given in  
 CC AAB43398 to AAB44239. The proteins can have activities based on the  
 CC tissues and cells the genes are expressed in. Example of activities  
 CC include: cytostatic; proliferative; vulnery; immunomodulatory;  
 CC antidiabetic; antiasthmatic; antirheumatic; antiarthritic;  
 CC antiinflammatory; antithyroid; antiallergic; antibacterial; antiviral;  
 CC dermatological; neuroprotective; cardiant; thrombolytic; coagulant;  
 CC nootropic; vasotropic; antipsoriatic and antiangiogenic. The  
 CC polynucleotides and polypeptides can be used for preventing, treating or  
 CC ameliorating medical conditions and diagnosing pathological conditions.  
 CC Polynucleotides, polypeptides, antibodies, agonists and antagonists from  
 CC the present invention may be used to treat immune disorders by activating  
 CC or inhibiting the proliferation, differentiation or mobilisation of  
 CC immune cells, to treat disorders of haematopoietic cells, autoimmune  
 CC disorders, allergic reactions, graft versus host disease, and organ  
 CC rejection, modulate haemostatic or thrombolytic activity, modulate  
 CC inflammation, cancers, cardiovascular disorders, neurological disease and  
 CC bacterial or viral infections. The peptides, nucleotides, antibodies,  
 CC agonists and antagonists may be also be used in drug screens. AAC78449 to  
 CC AAC78457 and AAB44240 represent sequences used in the exemplification of  
 CC the present invention  
 XX

```

SQ Sequence 131 AA;
  Query Match      100.0%; Score 597; DB 3; Length 131;
  Best Local Similarity 100.0%; Pred. No. 1.6e-60;
  Matches 115; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSGEGQTSVAPPPPEVEPGSGVRIIVVEYCEPGFATYLEASAVKEQYPGIEISRLG 60
Db |||||
Qy 17 MSGEGQTSVAPPPPEVEPGSGVRIIVVEYCEPGFATYLEASAVKEQYPGIEISRLG 76
Db |||||
Qy 61 GTGAPEIEINGQLVFSKLENGGFPYEKDLIEAIRRASNGETLEKITSRPPCVIL 115
Db |||||
Qy 77 GTGAPEIEINGQLVFSKLENGGFPYEKDLIEAIRRASNGETLEKITSRPPCVIL 131
Db |||||

RESULT 14
ADK49069
ID ADK49069 standard; protein; 131 AA.
XX
AC ADK49069;
XX
DT 04-NOV-2004 (first entry)
XX
DE Human breast/bladder carcinoma C35-related protein - SEQ ID 155.
XX
KW C35 epitope; cytostatic; vaccine; tumour; breast; bladder carcinoma;
KW human.
XX
OS Homo sapiens.
XX
PN WO2003104428-A2.
XX
PD 18-DEC-2003.
XX
PF 10-JUN-2003; 2003WO-US018252.
XX
PR 10-JUN-2002; 2002US-0386738P.
PR 11-DEC-2002; 2002US-0432241P.
PR 23-APR-2003; 2003US-0464650P.
XX
PA (VACC-) VACCINEX INC.
PA (UYRP) UNIV ROCHESTER.
XX
PI Zauderer M, Evans EE, Borrello MA;
XX
DR WPI; 2004-062349/06.
XX
PT Novel C35 polypeptide useful for formulation of immunogenic composition
PT to induce antibodies and cell-mediated immunity against tumor cells.
XX
PS Disclosure; SEQ ID NO 155; 626pp; English.
XX
CC The invention relates to a novel isolated polypeptide comprising or
CC consisting of two or more C35 peptide epitopes. The polypeptide of the
CC invention demonstrates cytostatic activity and may be useful for the
CC formulation of an immunogenic composition, such as a vaccine, to induce
CC antibodies and cell-mediated immunity against target cells such as tumour
CC cells. Furthermore, the polypeptide and its analogues may be useful as
CC prognostic markers for carcinoma, such as human breast or bladder
CC carcinoma. The current sequence is that of human breast/bladder carcinoma
CC C35-related protein of the invention.
XX
SQ Sequence 131 AA;
  Query Match      100.0%; Score 597; DB 8; Length 131;
  Best Local Similarity 100.0%; Pred. No. 1.6e-60;
  Matches 115; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSGEGQTSVAPPPPEVEPGSGVRIIVVEYCEPGFATYLEASAVKEQYPGIEISRLG 60
Db |||||
Qy 17 MSGEGQTSVAPPPPEVEPGSGVRIIVVEYCEPGFATYLEASAVKEQYPGIEISRLG 76
Db |||||
Qy 61 GTGAPEIEINGQLVFSKLENGGFPYEKDLIEAIRRASNGETLEKITSRPPCVIL 115
Db |||||

Db 77 GTGAPEIEINGQLVFSKLENGGFPYEKDLIEAIRRASNGETLEKITSRPPCVIL 131
|||
RESULT 15
AEA62539
ID AEA62539 standard; protein; 149 AA.
XX
AC AEA62539;
XX
DT 25-AUG-2005 (first entry)
XX
DE Recombinant C35 antigen.
XX
KW C35 antigen; antigen; cell growth; cancer; cytostatic; apoptosis;
KW immunotherapy; hyperproliferation; epitope mapping.
XX
OS Homo sapiens.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Peptide 1..29
FT Protein /note= "His-tagged signal peptide"
FT Protein 30..149
FT Protein /note= "C35 antigen"
XX
PN WO2005055936-A2.
XX
PD 23-JUN-2005.
XX
PF 06-DEC-2004; 2004WO-US040573.
XX
PR 04-DEC-2003; 2003US-0526572P.
PR 23-DEC-2003; 2003US-0531688P.
XX
PA (VACC-) VACCINEX INC.
XX
PI Evans EE, Paris MJ, Sahasrabudhe DM, Smith ES, Zauderer M;
XX
DR WPI; 2005-458501/46.
XX
PT Killing cancer cells, by administering apoptosis-inducing therapy and
PT administering antibody specific for intracellular, cancer-associated
PT protein other than C35, or antibody specific for C35.
XX
PS Example 11; Fig 9; 255pp; English.
XX
CC The invention relates to killing (M1) cancer cells, comprising
CC administering an apoptosis-inducing therapy to cancer cells, and
CC administering to the cells an antibody specific for an intracellular,
CC cancer-associated protein, provided that the protein is not C35 antigen,
CC where protein becomes exposed on the cell surface in cells undergoing
CC apoptosis, where the antibody is conjugated to or complexed with a toxin.
CC The non-C35 antigen protein is a prenylated protein. Also included are an
CC isolated antibody (I) specific for C35 (chosen from an antibody
CC comprising the VH region encoded by clone 1B3G, the VL region encoded by
CC clone 1B3K, the VH region encoded by clone 1P2G, the VL region encoded by
CC clone 1P2K, the VH region encoded by clone H0009, the VL region encoded
CC by clone L0010, an antibody comprising at least one of CDR1 or CDR2 of
CC the VH region encoded by AEA62495, an antibody comprising at least one of
CC CDR1 or CDR2 of the VH region encoded by AEA62499, an antibody comprising
CC at least one of CDR1, CDR2, or CDR3 of the VL region encoded by
CC AEA62497, a chimeric antibody, or a humanized antibody), a polynucleotide
CC encoding the antibody, a vector comprising the polynucleotide, a host
CC cell comprising the vector and a composition comprising the antibody and
CC preferably human in need of eradication of smaller tumors and/or
CC micrometastases, or in need of cancer treatment for C35-associated cancer
CC chosen from breast cancer, ovarian cancer, bladder cancer, lung cancer,
CC prostate cancer, pancreatic cancer, colon cancer, melanoma and other
CC hyperproliferative disorders. The antibody is useful for detecting,
CC diagnosing or monitoring C35-associated cancers. The antibody comprises a
CC chimeric antibody comprising human immunoglobulin constant regions fused

```

CC to the variable regions of mouse anti-35 antibodies (named 1F2, 1B3, MAb  
CC 165 and MAb 171). The present sequence represents a recombinant human C35  
CC antigen (with a used His-tagged signal peptide), Lys-C digests of which  
CC were used in epitope mapping studies.

XX  
SQ Sequence 149 AA;

Query Match 100.0%; Score 597; DB 9; Length 149;  
Best Local Similarity 100.0%; Pred. No. 1.9e-60;  
Matches 115; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 MSGEPGQTSVAPPPEVEPGSGVRIIVVEYCEPGFEATYLELASAVKEQYPGIEIESRLG 60  
|||  
Db 35 MSGEPGQTSVAPPPEVEPGSGVRIIVVEYCEPGFEATYLELASAVKEQYPGIEIESRLG 94  
|||  
  
QY 61 GTGAPEIEINGQLVFSKLENGGPPYEKDLIEAIRRASNGETLEKITNSRPPCVIL 115  
|||  
Db 95 GTGAPEIEINGQLVFSKLENGGPPYEKDLIEAIRRASNGETLEKITNSRPPCVIL 149  
|||

Search completed: December 6, 2006, 09:24:45  
Job time : 149 secs

**This Page Blank (uspto)**

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: December 6, 2006, 10:45:18 ; Search time 83 Seconds  
(without alignments)  
641.804 Million cell updates/sec

Title: US-09-824-787B-2

Perfect score: 597

Sequence: 1 MSCEPGQTSVAPPEEVEPG.....ASNGTLEKITNSRPPCVIL 115

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2097797 seqs, 463214858 residues

Total number of hits satisfying chosen parameters: 2097797

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications\_AA\_Main:\*

- 1: /EMC\_Celerra\_SID33/ptodata/2/pubpaa/US07\_PUBCOMB.pep:\*
- 2: /EMC\_Celerra\_SID33/ptodata/2/pubpaa/US08\_PUBCOMB.pep:\*
- 3: /EMC\_Celerra\_SID33/ptodata/2/pubpaa/US09\_PUBCOMB.pep:\*
- 4: /EMC\_Celerra\_SID33/ptodata/2/pubpaa/US10\_PUBCOMB.pep:\*
- 5: /EMC\_Celerra\_SID33/ptodata/2/pubpaa/US10B\_PUBCOMB.pep:\*
- 6: /EMC\_Celerra\_SID33/ptodata/2/pubpaa/US11\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	597	100.0	115	3	US-09-824-787B-2
2	597	100.0	115	4	US-10-435-696-93
3	597	100.0	115	4	US-10-457-829-2
4	597	100.0	115	5	US-10-887-230-2
5	597	100.0	115	5	US-10-855-988-86
6	597	100.0	115	6	US-11-003-819-2
7	597	100.0	115	6	US-11-155-288-11
8	597	100.0	115	6	US-11-233-510-6
9	597	100.0	117	3	US-09-833-203-34
10	597	100.0	124	4	US-10-264-049-4187
11	597	100.0	131	3	US-09-925-301-966
12	597	100.0	131	4	US-10-457-829-155
13	597	100.0	206	4	US-10-177-293-480
14	403	67.5	90	4	US-10-457-829-153
15	148.5	24.9	95	6	US-11-097-143-38691
16	129	21.6	250	4	US-10-424-599-153780
17	126.5	21.2	244	5	US-10-739-930-10642
18	121	20.3	237	4	US-10-767-701-41097
19	119	19.9	232	4	US-10-437-963-106912
20	114	19.1	257	4	US-10-437-963-181063
21	113	18.9	250	4	US-10-767-701-44890
22	109	18.3	250	4	US-10-425-115-229236
23	108.5	18.2	228	4	US-10-424-559-149832
24	108.5	18.2	257	4	US-10-425-115-212160
25	87	14.6	198	6	US-11-097-143-7521
26	85.5	14.3	74	4	US-10-264-049-2688
27	80	13.4	137	3	US-09-374-046A-96

28	80	13.4	137	4	US-10-616-263-96	Sequence 96, Appl
29	80	13.4	146	3	US-09-892-877-301	Sequence 301, App
30	80	13.4	146	3	US-09-948-783-314	Sequence 314, App
31	80	13.4	163	4	US-10-614-853-14	Sequence 14, Appl
32	79	13.2	469	4	US-10-369-493-22836	Sequence 22836, A
33	79	13.2	2042	4	US-10-437-963-112456	Sequence 112456, A
34	77	12.9	370	4	US-10-369-493-12313	Sequence 12313, A
35	76.5	12.8	170	6	US-11-096-568A-10110	Sequence 10110, A
36	76.5	12.8	186	6	US-11-096-568A-10109	Sequence 10109, A
37	76.5	12.8	198	6	US-11-096-568A-10108	Sequence 10108, A
38	76	12.7	282	4	US-10-437-963-136117	Sequence 136117, A
39	76	12.7	283	6	US-11-096-568A-31233	Sequence 31233, A
40	76	12.7	390	6	US-11-096-568A-31232	Sequence 31232, A
41	76	12.7	395	6	US-11-096-568A-31231	Sequence 31231, A
42	76	12.7	1463	4	US-10-437-963-163410	Sequence 163410, A
43	75.5	12.6	393	4	US-10-369-493-1082	Sequence 1082, Ap
44	75.5	12.6	590	4	US-10-437-963-114110	Sequence 114110, A
45	75	12.6	239	5	US-10-506-454-642	Sequence 642, App

#### ALIGNMENTS

##### RESULT 1

US-09-824-787B-2

; Sequence 2, Application US/09824787B

; Patent No. US20020155447A1

; GENERAL INFORMATION:

; APPLICANT: Zauderer, Maurice

; APPLICANT: Evans, Elizabeth E.

; APPLICANT: Borrello, Melinda A.

; TITLE OF INVENTION: A Gene Differentially Expressed in Breast and

; TITLE OF INVENTION: Bladder Cancer, and Encoded Polypeptides

; FILE REFERENCE: 1821.0040001

; CURRENT APPLICATION NUMBER: US/09/824,787B

; PRIOR FILING DATE: 2001-04-04

; PRIOR FILING DATE: 60/194,463

; NUMBER OF SEQ ID NOS: 147

; SOFTWARE: Patentin Ver. 2.1

; SEQ ID NO 2

; LENGTH: 115

; TYPE: PRT

; ORGANISM: Homo sapiens

; US-09-824-787B-2

Query Match 100.0%; Score 597; DB 3; Length 115;

Best Local Similarity 100.0%; Pred.No. 2.3e-58;

Matches 115; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSCEPGQTSVAPPEEVEPGSGVRIVVVEYCEPCGFEATYLELASAVKEQVPGIEISRLG 60

Db 1 MSCEPGQTSVAPPEEVEPGSGVRIVVVEYCEPCGFEATYLELASAVKEQVPGIEISRLG 60

Qy 61 GTGAFIEINGQLVFSKLENGGFPFYKDLIEAIRRASNGTLEKITNSRPPCVIL 115

Db 61 GTGAFIEINGQLVFSKLENGGFPFYKDLIEAIRRASNGTLEKITNSRPPCVIL 115

##### RESULT 2

US-10-435-696-93

; Sequence 93, Application US/10435696

; Publication No. US20040018525A1

; GENERAL INFORMATION:

; APPLICANT: Wirtz, Ralph

; APPLICANT: Munnes, Marc

; APPLICANT: Kallabis, Harald

; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE PREDICTION, DIAGNOSIS, PROGNOSIS

; TITLE OF INVENTION: PREVENTION AND TREATMENT OF MALIGNANT NEOPLASIA

; FILE REFERENCE: Lea 36 108

; CURRENT APPLICATION NUMBER: US/10/435,696

; PRIOR FILING DATE: 2003-05-09

; PRIOR APPLICATION NUMBER: EP03003112.4





US-11-003-819-2  
; Sequence 2, Application US/11003819  
; Publication No. US20050158323A1  
; GENERAL INFORMATION:  
; APPLICANT: Evans, Elizabeth E.  
; APPLICANT: Paris, Mark J.  
; APPLICANT: Sahasrabudhe, Deepak M.  
; APPLICANT: Zauderer, Maurice  
; APPLICANT: Smith, Ernest S.  
; TITLE OF INVENTION: Methods of Killing Tumor Cells by Targeting Internal Antigens  
; TITLE OF INVENTION: Exposed on Apoptotic Tumor Cells  
; FILE REFERENCE: 1843.0190002  
; CURRENT APPLICATION NUMBER: US/11/003,819  
; CURRENT FILING DATE: 2004-12-06  
; PRIOR APPLICATION NUMBER: US 60/256,572  
; PRIOR FILING DATE: 2003-12-04  
; PRIOR APPLICATION NUMBER: US 60/531,688  
; PRIOR FILING DATE: 2003-12-23  
; NUMBER OF SEQ ID NOS: 61  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 2  
; LENGTH: 115  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-11-003-819-2

Query Match 100.0%; Score 597; DB 6; Length 115;  
Best Local Similarity 100.0%; Pred. No. 2.3e-58;  
Matches 115; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MSSEPGQTSVAPPPPEEVEPGSGVRIIVVEYCEPCGFEATYLELASAVKEQYPGIEISRLG 60  
Db 1 MSSEPGQTSVAPPPPEEVEPGSGVRIIVVEYCEPCGFEATYLELASAVKEQYPGIEISRLG 60  
QY 61 GTGAFIEINGQLVFSKLENGGPPYEKDLIEAIRRASNGTLEKITSRRPPCVIL 115  
Db 61 GTGAFIEINGQLVFSKLENGGPPYEKDLIEAIRRASNGTLEKITSRRPPCVIL 115

RESULT 7  
US-11-155-288-11  
; Sequence 11, Application US/11155288  
; Publication No. US20060008468A1  
; GENERAL INFORMATION:  
; APPLICANT: Chiang, Chih-Sheng  
; APPLICANT: Simard, John J.D.  
; TITLE OF INVENTION: COMBINATIONS OF TUMOR-ASSOCIATED  
; TITLE OF INVENTION: ANTIGENS IN DIAGNOSTICS FOR VARIOUS TYPES OF CANCERS  
; FILE REFERENCE: MANNK.050A  
; CURRENT APPLICATION NUMBER: US/11/155,288  
; CURRENT FILING DATE: 2005-06-17  
; PRIOR APPLICATION NUMBER: 60/580,969  
; PRIOR FILING DATE: 2004-06-17  
; NUMBER OF SEQ ID NOS: 40  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 11  
; LENGTH: 115  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-11-155-288-11

Query Match 100.0%; Score 597; DB 6; Length 115;  
Best Local Similarity 100.0%; Pred. No. 2.3e-58;  
Matches 115; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MSSEPGQTSVAPPPPEEVEPGSGVRIIVVEYCEPCGFEATYLELASAVKEQYPGIEISRLG 60  
Db 1 MSSEPGQTSVAPPPPEEVEPGSGVRIIVVEYCEPCGFEATYLELASAVKEQYPGIEISRLG 60  
QY 61 GTGAFIEINGQLVFSKLENGGPPYEKDLIEAIRRASNGTLEKITSRRPPCVIL 115  
Db 61 GTGAFIEINGQLVFSKLENGGPPYEKDLIEAIRRASNGTLEKITSRRPPCVIL 115

RESULT 8  
US-11-233-510-6  
; Sequence 6, Application US/11233510  
; Publication No. US20060063190A1  
; GENERAL INFORMATION:  
; APPLICANT: Fischer, Timothy J.  
; APPLICANT: Whitehead, Clark M.  
; APPLICANT: Malinowski, Douglas P.  
; APPLICANT: Marcelloni, Raphael  
; APPLICANT: Morel, Didier  
; TITLE OF INVENTION: Methods and Compositions for Evaluating  
; TITLE OF INVENTION: Breast Cancer Prognosis  
; FILE REFERENCE: 46143/296738  
; CURRENT APPLICATION NUMBER: US/11/233,510  
; CURRENT FILING DATE: 2005-09-22  
; PRIOR APPLICATION NUMBER: 60/612,073  
; PRIOR FILING DATE: 2004-09-22  
; PRIOR APPLICATION NUMBER: 60/611,965  
; PRIOR FILING DATE: 2004-09-22  
; NUMBER OF SEQ ID NOS: 41  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 6  
; LENGTH: 115  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-11-233-510-6

Query Match 100.0%; Score 597; DB 6; Length 115;  
Best Local Similarity 100.0%; Pred. No. 2.3e-58;  
Matches 115; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MSSEPGQTSVAPPPPEEVEPGSGVRIIVVEYCEPCGFEATYLELASAVKEQYPGIEISRLG 60  
Db 1 MSSEPGQTSVAPPPPEEVEPGSGVRIIVVEYCEPCGFEATYLELASAVKEQYPGIEISRLG 60  
QY 61 GTGAFIEINGQLVFSKLENGGPPYEKDLIEAIRRASNGTLEKITSRRPPCVIL 115  
Db 61 GTGAFIEINGQLVFSKLENGGPPYEKDLIEAIRRASNGTLEKITSRRPPCVIL 115

RESULT 9  
US-09-833-203-34  
; Sequence 34, Application US/09833203  
; Publication No. US20030166277A1  
; GENERAL INFORMATION:  
; APPLICANT: Zauderer, Maurice  
; APPLICANT: Smith, Ernest S.  
; TITLE OF INVENTION: Targeted Vaccine Delivery Systems  
; FILE REFERENCE: 1821.0020001  
; CURRENT APPLICATION NUMBER: US/09/833,203  
; CURRENT FILING DATE: 2001-04-12  
; PRIOR APPLICATION NUMBER: US 60/196,472  
; PRIOR FILING DATE: 2000-04-12  
; NUMBER OF SEQ ID NOS: 63  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 34  
; LENGTH: 117  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: C35  
US-09-833-203-34

Query Match 100.0%; Score 597; DB 3; Length 117;  
Best Local Similarity 100.0%; Pred. No. 2.3e-58;  
Matches 115; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MSSEPGQTSVAPPPPEEVEPGSGVRIIVVEYCEPCGFEATYLELASAVKEQYPGIEISRLG 60  
Db 3 MSSEPGQTSVAPPPPEEVEPGSGVRIIVVEYCEPCGFEATYLELASAVKEQYPGIEISRLG 62

```
QY 61 GTGAFIEINGQLVFSKLENGGPPYKDLIEAIRRASNGETLEKITSRPPCVIL 115
DB 63 GTGAFIEINGQLVFSKLENGGPPYKDLIEAIRRASNGETLEKITSRPPCVIL 117

RESULT 10
US-10-264-049-4187
; Sequence 4187, Application US/10264049
; Publication No. US20040005579A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA133PI
; CURRENT APPLICATION NUMBER: US/10/264,049
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/18569
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: US 60/209,467
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 4360
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 4187
; LENGTH: 124
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-264-049-4187

Query Match 100.0%; Score 597; DB 4; Length 124;
Best Local Similarity 100.0%; Pred. No. 2.5e-58;
Matches 115; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSGEPQQTSAVAPPPEVEPGSGVRIIVVEYCEPCGPEATYLELASAVKEQYPGIEISRLG 60
DB 10 MSGEPQQTSAVAPPPEVEPGSGVRIIVVEYCEPCGPEATYLELASAVKEQYPGIEISRLG 69
QY 61 GTGAFIEINGQLVFSKLENGGPPYKDLIEAIRRASNGETLEKITSRPPCVIL 115
DB 70 GTGAFIEINGQLVFSKLENGGPPYKDLIEAIRRASNGETLEKITSRPPCVIL 124

RESULT 11
US-09-925-301-966
; Sequence 966, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 966
; LENGTH: 131
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-301-966

Query Match 100.0%; Score 597; DB 3; Length 131;
Best Local Similarity 100.0%; Pred. No. 2.7e-58;
Matches 115; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSGEPQQTSAVAPPPEVEPGSGVRIIVVEYCEPCGPEATYLELASAVKEQYPGIEISRLG 60
DB 17 MSGEPQQTSAVAPPPEVEPGSGVRIIVVEYCEPCGPEATYLELASAVKEQYPGIEISRLG 76
QY 61 GTGAFIEINGQLVFSKLENGGPPYKDLIEAIRRASNGETLEKITSRPPCVIL 115
DB 77 GTGAFIEINGQLVFSKLENGGPPYKDLIEAIRRASNGETLEKITSRPPCVIL 131

QY 61 GTGAFIEINGQLVFSKLENGGPPYKDLIEAIRRASNGETLEKITSRPPCVIL 115
DB 63 GTGAFIEINGQLVFSKLENGGPPYKDLIEAIRRASNGETLEKITSRPPCVIL 117
```

```
RESULT 12
US-10-457-829-155
; Sequence 155, Application US/10457829
; Publication No. US20040063907A1
; GENERAL INFORMATION:
; APPLICANT: Zauderer, Maurice
; APPLICANT: Evans, Elizabeth S.
; APPLICANT: Borrello, Melinda A.
; TITLE OF INVENTION: A Gene Differentially Expressed in Breast and
; TITLE OF INVENTION: Bladder Cancer, and Encoded Polypeptides
; FILE REFERENCE: 1821.0040005
; CURRENT APPLICATION NUMBER: US/10/457,829
; CURRENT FILING DATE: 2003-06-10
; PRIOR APPLICATION NUMBER: US 60/464,650
; PRIOR FILING DATE: 2003-04-23
; NUMBER OF SEQ ID NOS: 160
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 155
; LENGTH: 131
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-457-829-155

Query Match 100.0%; Score 597; DB 4; Length 131;
Best Local Similarity 100.0%; Pred. No. 2.7e-58;
Matches 115; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSGEPQQTSAVAPPPEVEPGSGVRIIVVEYCEPCGPEATYLELASAVKEQYPGIEISRLG 60
DB 17 MSGEPQQTSAVAPPPEVEPGSGVRIIVVEYCEPCGPEATYLELASAVKEQYPGIEISRLG 76
QY 61 GTGAFIEINGQLVFSKLENGGPPYKDLIEAIRRASNGETLEKITSRPPCVIL 115
DB 77 GTGAFIEINGQLVFSKLENGGPPYKDLIEAIRRASNGETLEKITSRPPCVIL 131

RESULT 13
US-10-177-293-480
; Sequence 480, Application US/10177293
; Publication No. US20030124128A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Glatt, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Gannavartu, Manjula
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Mertens, Maureen
; APPLICANT: Myer, Vic
; APPLICANT: Wang, Youzhen
; APPLICANT: Xu, Yongyao
; APPLICANT: Hoerach, Sebastian
; APPLICANT: Monahan, John
; APPLICANT: Meyers, Rachel E.
; APPLICANT: East Jr., Robert C.
; APPLICANT: Hortobagyi, Gabriel N.
; APPLICANT: Pusztai, Lajos
; APPLICANT: Meric, Funda
; APPLICANT: Sahin, Aysegul
; APPLICANT: Mills, Gordon B.
; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT,
; TITLE OF INVENTION: PREVENTION, AND THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-038
; CURRENT APPLICATION NUMBER: US/10/177,293
; CURRENT FILING DATE: 2002-06-21
; PRIOR APPLICATION NUMBER: US 60/299,887
; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: US 60/301,572
; PRIOR FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: US 60/306,501
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: US 60/325,002
```

;; PRIOR FILING DATE: 2001-09-25  
;; PRIOR APPLICATION NUMBER: US 60/362,585  
;; PRIOR FILING DATE: 2002-03-05  
;; PRIOR APPLICATION NUMBER: US 60/xxx,xxx  
;; PRIOR FILING DATE: 2002-05-14  
;; NUMBER OF SEQ ID NOS: 506  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 480  
;; LENGTH: 206  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-10-177-293-480

Query Match 100.0%; Score 597; DB 4; Length 206;  
Best Local Similarity 100.0%; Pred. No. 5e-58;  
Matches 115; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MSCEPGQTSVAPPEEVEPGSGVRIIVVEYCEPCGFEATYLELASAVKEQYPGIEISRLG 60  
Db 92 MSCEPGQTSVAPPEEVEPGSGVRIIVVEYCEPCGFEATYLELASAVKEQYPGIEISRLG 151  
QY 61 GTGAFIEINGQLVFSKLENGGPPYEKDLIEAIRRASNGETLEKITNSRPPCVIL 115  
Db 152 GTGAFIEINGQLVFSKLENGGPPYEKDLIEAIRRASNGETLEKITNSRPPCVIL 206

RESULT 14  
US-10-457-829-153  
;; Sequence 153, Application US/10457829  
;; Publication No. US20040063907A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Zauderer, Maurice  
;; APPLICANT: Evans, Elizabeth E.  
;; APPLICANT: Borrello, Melinda A.  
;; TITLE OF INVENTION: A Gene Differentially Expressed in Breast and  
;; TITLE OF INVENTION: Bladder Cancer, and Encoded Polypeptides  
;; FILE REFERENCE: 1821.0040005  
;; CURRENT APPLICATION NUMBER: US/10/457,829  
;; PRIOR FILING DATE: 2003-06-10  
;; PRIOR APPLICATION NUMBER: US 60/464,650  
;; PRIOR FILING DATE: 2003-04-23  
;; NUMBER OF SEQ ID NOS: 160  
;; SOFTWARE: PatentIn Ver. 2.1  
;; SEQ ID NO 153  
;; LENGTH: 90  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
;; FEATURE:  
;; NAME/KEY: MISC FEATURE  
;; LOCATION: (18)..(18)  
;; OTHER INFORMATION: Xaa is an unknown amino acid  
;; FEATURE:  
;; NAME/KEY: MISC FEATURE  
;; LOCATION: (46)..(46)  
;; OTHER INFORMATION: Xaa is an unknown amino acid  
;; FEATURE:  
;; NAME/KEY: MISC FEATURE  
;; LOCATION: (78)..(78)  
;; OTHER INFORMATION: Xaa is an unknown amino acid  
US-10-457-829-153

Query Match 67.5%; Score 403; DB 4; Length 90;  
Best Local Similarity 87.6%; Pred. No. 7.3e-37;  
Matches 78; Conservative 2; Mismatches 9; Indels 0; Gaps 0;  
QY 1 MSCEPGQTSVAPPEEVEPGSGVRIIVVEYCEPCGFEATYLELASAVKEQYPGIEISRLG 60  
Db 1 MSCEPGQTSVAPPEEVEPGSGVRIIVVEYCEPCGFEATYLELASAVKEQYPGIEISRLG 60  
QY 61 GTGAFIEINGQLVFSKLENGGPPYEKDL 89  
Db 61 GTGAFIEINGQLVFSKLENGGPPYEKDV 89

RESULT 15  
US-11-097-143-38691  
;; Sequence 38691, Application US/11097143  
;; Publication No. US20050208558A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Venter, J. Craig  
;; APPLICANT: et al.  
;; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID  
;; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE  
;; TITLE OF INVENTION: DROSOPHILA GENES.  
;; FILE REFERENCE: CL000728  
;; CURRENT APPLICATION NUMBER: US/11/097,143  
;; PRIOR FILING DATE: 2005-04-04  
;; PRIOR APPLICATION NUMBER: 60/157,832  
;; PRIOR FILING DATE: 1999-10-05  
;; PRIOR APPLICATION NUMBER: 60/160,191  
;; PRIOR FILING DATE: 1999-10-19  
;; PRIOR APPLICATION NUMBER: 60/161,932  
;; PRIOR FILING DATE: 1999-10-28  
;; PRIOR APPLICATION NUMBER: 60/164,769  
;; PRIOR FILING DATE: 1999-11-12  
;; PRIOR APPLICATION NUMBER: 60/173,383  
;; PRIOR FILING DATE: 1999-12-28  
;; PRIOR APPLICATION NUMBER: 60/175,693  
;; PRIOR FILING DATE: 2000-01-12  
;; PRIOR APPLICATION NUMBER: 60/184,831  
;; PRIOR FILING DATE: 2000-02-24  
;; PRIOR APPLICATION NUMBER: 60/191,637  
;; PRIOR FILING DATE: 2000-03-23  
;; NUMBER OF SEQ ID NOS: 43008  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 38691  
;; LENGTH: 95  
;; TYPE: PRT  
;; ORGANISM: DROSOPHILA  
US-11-097-143-38691

Query Match 24.9%; Score 148.5; DB 6; Length 95;  
Best Local Similarity 34.0%; Pred. No. 2e-08;  
Matches 32; Conservative 17; Mismatches 44; Indels 1; Gaps 1;  
QY 23 VRIVVEYCEPCGFEATYLELASAVKEQYPGIEISRLGTTGAFIEINGQLVFSKLENGG 82  
Db 2 VRVEYCGICNFGQCHLLREFLLASSPDLDISCTGRRGSEVSDGQLVHSLSLA 61  
QY 83 FPYEKDLIEAIRRASNGETLEKITNSR-PPCVIL 115  
Db 62 FPOHASVLAQVQKAERGEPEVKVLEQPIKDCVVM 95

Search completed: December 6, 2006, 10:50:07  
Job time : 85 secs

**This Page Blank (uspto)**